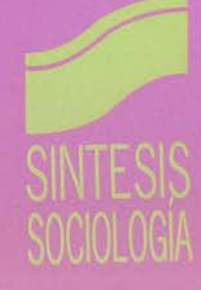
METODOLOGÍA CUANTITATIVA
Estrategias y técnicas de investigación social

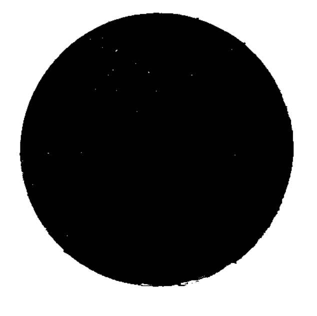
M.ª Ángeles Cea D'Ancona





SINTESIS
SOCIOLOGÍA

METODOLOGÍA CUANTITATIVA: ESTRATEGIAS Y TÉCNICAS DE INVESTIGACIÓN SOCIAL



PROYECTO EDITORIAL
SÍNTESIS SOCIOLOGÍA

Directores:
Antonio Izquierdo Escribano
Jesús Leal Maldonado
Ramón Ramos Torre

METODOLOGÍA CUANTITATIVA: ESTRATEGIAS Y TÉCNICAS DE INVESTIGACIÓN SOCIAL

M.ª Ángeles Cea D'Ancona

Diseño de cubierta: JV Diseño gráfico

Depósito Legal: M. 30.037-1996
ISBN: 84-7738-420-7

Impresión: Lavel, S. A.
Impreso en España - Printed in Spain

Al lector anónimo

AGRADECIMIENTOS

Agradezco al director de la colección, Antonio Izquierdo, y al editor, Francisco Belloso, el haberme propuesto la elaboración del presente manual.

También, deseo dejar constancia de mi gratitud hacia los profesores y autores que han contribuido a mi formación, la mayoría de cuyos nombres aparecen listados en la bibliografía. Mención especial merece Miguel S. Valles Martínez, por la lectura reposada de borradores y la aportación de su punto de vista personal y profesional.

A la Fundación Juan March, y a otros muchos centros, agradezco el haberme facilitado la consulta de sus fondos bibliográficos.

Asimismo, tengo presente a las promociones de alumnos de los últimos años, por su estímulo constante.

Mi último pensamiento se dirige a mi familia, a la que tanto debo.

ÍNDICE

SEGUNDA PARTE
LOS PRELIMINARES DE LA INVESTIGACIÓN SOCIAL

CAPÍTULO 3: LA ORGANIZACIÓN DE LA INVESTIGACIÓN

CAPÍTULO 4: LA OPERACIONALIZACIÓN DE CONCEPTOS

CAPÍTULO 5: LA SELECCIÓN DE LAS UNIDADES DE OBSERVACIÓN: EL DISEÑO DE LA MUESTRA

TERCERA PARTE
LA OBTENCIÓN DE INFORMACIÓN

CAPÍTULO 6: EL USO DE FUENTES DOCUMENTALES Y ESTADÍSTICAS

CUARTA PARTE

EL ANÁLISIS Y LA PRESENTACIÓN DE LA INFORMACIÓN

CAPÍTULO 9: EL ANÁLISIS DE LOS DATOS

CAPÍTULO 10: EL ANÁLISIS DE CONTENIDO CUANTITATIVO

INTRODUCCIÓN

El texto que el lector tiene en sus manos pretende ser un *"manual didáctico"*, que despierte el interés por la *metodología* de la investigación, a la vez que instruya en los conocimientos necesarios para su práctica.

La traducción de este *desideratum* se ha plasmado en la escritura de once capítulos agrupados en cuatro bloques temáticos, de los que se da pronta cuenta en el índice del texto. En ellos han adquirido forma los contenidos derivados de los tres conceptos que componen el título y la clave de la obra: *"metodología"*, *"estrategias"* y *"técnicas"* de investigación.

Estos conceptos se entienden de manera inclusiva: el primero incluyendo al segundo y éste, a su vez, al término más subordinado, más coloquial también, de *técnicas*. La *metodología* ocupa el orden superior (supraordinado). Comprende los "principios generales que guían la investigación" (Bulmer, 1992: 4). Del nivel metodológico se pasa al nivel estratégico. En él se deciden las *técnicas* concretas para la consecución de los objetivos del estudio, en consonancia con el enfoque que quiera darse a la indagación.

La exposición de la materia comienza con la presentación de la génesis y el desarrollo de la sociología como ciencia para, desde la historia, comprender el pluralismo teórico-metodológico existente en este campo y la necesidad de convergencia o articulación interparadigmática. De ello resulta una propuesta de definición y estructuración de la investigación social.

A continuación se presenta un mapa organizativo (cognitivo), en el que se ponen en relación las *técnicas* con las *estrategias,* y éstas con los *diseños*. Después, se desarrolla cada fase del proceso de investigación.

El segundo bloque temático se dedica a "los preliminares de la investigación social". Incluye las fases iniciales del *proyecto de investigación:* desde la formulación del

problema de estudio, hasta su operacionalización en la vertiente teórico-analítica (conceptos y variables a analizar) y poblacional (el diseño de la muestra).

Siguiendo el orden del proceso de investigación, la tercera parte se ocupa de la obtención de información en la *metodología cuantitativa* mediante las principales *estrategias* y *técnicas* de investigación. De cada una de ellas se señalan sus características esenciales, junto con las ventajas e inconvenientes fundamentales de su uso en la investigación social.

El libro concluye con "el análisis y la presentación de la información". En este último bloque temático se ofrece una visión panorámica de las distintas técnicas estadísticas de análisis, así como una breve relación de pautas a seguir en la redacción del Informe de investigación.

La extensión de los capítulos es muy dispar. Depende de la complejidad de los contenidos temáticos correspondientes. Para facilitar la comprensión de éstos, la exposición teórica se acompaña de varios *ejemplos* ilustrativos. A estos se suma la descripción metodológica de investigaciones "reales", efectuadas en España, en años recientes (como ejemplos desarrollados). Con la presentación de estudios realizados en un contexto próximo al del lector se quiere facilitar la comprensión del ejercicio de la investigación social, además de ofrecer referencias que pueden serle útiles en su actividad profesional.

Los capítulos concluyen con una selección bibliográfica, a modo de "lecturas complementarias" recomendadas, y una propuesta de "ejercicios". Al enfrentarse a su realización, el lector podrá comprobar si ha asimilado los contenidos esenciales de cada capítulo.

Si al final de la lectura del texto, se es capaz de *diseñar* una investigación adecuada a un problema específico de estudio, y se es capaz de justificar por qué se han seleccionado unas *estrategias* y *técnicas* con preferencia a otras, se habrán logrado los objetivos propuestos en la redacción de este manual.

PRIMERA PARTE

FUNDAMENTOS TEÓRICO-EMPÍRICOS DE LA INVESTIGACIÓN SOCIAL

1

COORDENADAS HISTÓRICAS DE LA INVESTIGACIÓN SOCIAL: TEORÍA E INVESTIGACIÓN EMPÍRICA

Antes de pasar a desarrollar los contenidos concretos de la materia que da título a este libro, resulta de interés remontarnos a algunos *antecedentes históricos:* a las diferentes aportaciones que desde la *teoría* y la *investigación empírica* han contribuido, en general, al desarrollo de la sociología y, más concretamente, a la configuración de los *métodos* y *técnicas de investigación* social. Este repaso histórico sin duda ayudará a una mejor comprensión y contextualización de la expresión *investigación social*. Como Rodríguez Ibáñez señala (1989: 22), "si la historia aparece es porque la discusión acerca de la teoría y del método no puede prescindir de ella".

1.1. Bases empíricas y teórico metodológicas en el origen de la sociología

Trazar los orígenes de cualquier disciplina científica siempre resulta polémico y sujeto a continuas revisiones. Depende de las distintas aportaciones que los estudiosos de la historia van haciendo. La sociología no es una excepción. Sus antecedentes varían de un autor a otro. Para algunos la sociología, en tanto ciencia independiente que analiza los fenómenos sociales, nace en 1839, cuando Comte acuña el término *sociología.* Otros (como Caplow, 1977; Bottomore, 1978; García Ferrando, 1979; o Kovaliova, 1989) se remontan más allá. A las ideas o doctrinas sociopolíticas de los considerados *precursores* (Montesquieu, Rousseau, Saint-Simon, Hegel). A ellos suman otra fuente originaria de la sociología moderna: las *investigaciones empíricas* llevadas a cabo por estudiosos de formación diversa (médicos, hombres de negocios, ingenieros, biólogos o naturalistas). Estos investigadores estaban más interesados en describir los graves problemas que trajo consigo la revolución industrial (y el nacimiento del capitalismo), que en las cuestiones teórico-metodológicas.

No obstante, en el estudio de los orígenes y desarrollo de la sociología generalmente ha primado más la formación de la *teoría sociológica* que las *indagaciones empíricas*. Aunque, en realidad, ambas contribuyeron a su surgimiento y consolidación.

Como la historia de la *investigación social empírica* antecede a la *sociología académica,* este capítulo comienza por la descripción de los orígenes de la *investigación social.*

1.1.1. Orígenes de la investigación social

Los *antecedentes* inmediatos de la investigación social empírica suelen remontarse a los siglos XVII y XVIII, al movimiento denominado *estadística social*: un grupo de estudiosos interesados en la recogida y organización de datos económicos, demográficos y sociales. A diferencia de sus predecesores, se caracterizaron por promover la aplicación de los mismos procedimientos de medición utilizados en las ciencias naturales al estudio de los fenómenos sociales. Ésta fue su aportación fundamental: proporcionar una mayor precisión a la descripción de la sociedad, haciendo uso de los avances habidos, hasta la fecha, en el campo de la estadística.

De los tiempos del emperador Yao en China y, con posterioridad, a las civilizaciones asiria, egipcia, griega y romana, datan los primeros recuentos "censales". Su finalidad era, esencialmente, tributaria y militar: los estados necesitaban conocer los recursos económicos y humanos de los que disponían (para fines bélicos y de fiscalidad). Pero, es en el siglo XVII, cuando estas "cuantificaciones" adquieran carácter científico. Ello fue posible gracias a las aportaciones fundamentales de una serie de autores englobados en dos *escuelas estadísticas* principales: los aritméticos políticos ingleses y la escuela estadística alemana.

a) Los aritméticos políticos ingleses

Una corriente pragmática constituida por científicos naturales ingleses, entre los cuales destacaron Graunt, Petty, Davenant, King, y Halley. Estos estudiosos buscaban la aplicación de sus conocimientos sobre *medición* al estudio de las actividades sociopolíticas. La cuantificación de las *regularidades sociales* debía hacerse a modo de las *regularidades naturales*.

En 1660, Petty acuña el término *"aritmética política"* en su obra *Political arithmetic.* En esta obra se insta a la *medición* de los fenómenos sociales, con los mismos instrumentos empleados en las ciencias naturales. Además, se defiende la utilización gubernamental de los datos estadísticos que se obtengan.

Graunt fue el autor que introduce, por vez primera, las *tablas de esperanza de vida*, para personas de diferentes grupos de edad. En ellas observa variaciones importantes (en pautas de mortalidad), entre los habitantes de entornos rurales y urbanos. Con posterioridad, estas *tablas* fueron ampliadas por el astrónomo Halley.

Por último, es también de resaltar la aportación de King al desarrollo de la *demografía* moderna. Principalmente, la clasificación que este investigador realiza de la población inglesa en 1688. Parte, para ello, de diferentes variables socioeconómicas, como el nivel de ingresos o el tamaño de la familia.

b) La escuela estadística alemana

También conocida como la *escuela estadística universitaria,* por su lugar de ubicación: universidades alemanas, como la de Berlín y Götinger.

A diferencia de la escuela anterior, la escuela alemana se caracteriza por la promoción de la denominada *geografía política.* En concreto, la *comparación* de datos demográficos, socioeconómicos y políticos, relativos a diferentes países.

Entre sus representantes destacan: Seckendorff (su promotor) en el siglo XVII; y, Conring (su fundador y representante principal), en el siglo XVIII, junto con Achenwall.

De ellos, Conring fue quien llevó a cabo más estudios comparativos. En ellos hacía uso de sistemas clasificatorios para la descripción y comparación de distintos países. Precisamente España fue uno de los países del que obtuvo un mayor número de datos sociopolíticos.

• En el siglo XVIII aparecen las primeras *investigaciones sociales empíricas,* propiamente dichas, de la mano de los llamados *reformadores.* Un conjunto de profesionales, de distintas disciplinas académicas (médicos, ingenieros, naturalistas), unidos por un mismo propósito: aportar evidencia empírica de los problemas sociales existentes en su época. Su finalidad era eminentemente política: movilizar a la opinión pública para exigir medidas concretas de *reforma social.* De ahí les viene el nombre.

Estos investigadores constituyeron (en sus países respectivos) *sociedades estadísticas,* que promovieron el desarrollo de las *encuestas sociales* en la investigación social. Los sistemas de recogida de información a través de parroquias y de las inspecciones fiscales (que hasta entonces habían imperado), se muestran ahora insuficientes para aportar la información que la situación social del momento demandaba.

Desde comienzos del siglo XIX, los distintos gobiernos fueron creando organismos oficiales para la realización de los *censos de población:* unas encuestas generales que sirvieran de registro de la distribución y condiciones de vida de la población en cada país. También se precisaba conocer la situación en la agricultura y en la industria.

De 1801 data el primer *censo de población* efectuado en Inglaterra y en Francia. Pero, mientras que en Inglaterra estas "encuestas generales" prosiguieron con una periodicidad de diez años, el gobierno francés las suprime en su país hasta el año 1836.

En España el primer *censo de población* (considerado oficial y moderno) aparece en 1857. A éste le siguieron los *censos* de 1860, 1877, 1887, 1897. Desde 1900 hasta 1970 se realizaron *censos de población* en los años acabados en cero (con referencia al 31 de diciembre). Los *censos* de 1981 y 1991 (con referencia a 1 de marzo) han supuesto la adecuación de las fechas censales españolas a la normativa europea.

• Respecto a las *encuestas sociales*, de las realizadas en Gran Bretaña destacan las siguientes:

a) El informe estadístico de Escocia (1791-1825), de John Sinclair

Una monumental investigación (en 21 volúmenes), llevada a cabo por John Sinclair, sobre los problemas de la población rural en Escocia. La información se obtuvo mediante la aplicación de un cuestionario *ex professo*. Este *cuestionario* incluía 116 preguntas referentes a diversas características sociodemográficas de la población (sexo, edad, ocupación, mortalidad, práctica religiosa), y de su entorno (orografía, recursos minerales y agrícolas). Para su realización, Sinclair contó con la colaboración del clero escocés, que le proporcionó información referente a 881 parroquias de Escocia.

Esta investigación tuvo una gran repercusión en numerosos países europeos. Concretamente, en la organización de los *censos* y en la elaboración de los *cuestionarios*.

b) Las condiciones morales y físicas de la vida de los obreros de la industria textil de Manchester (1832), de James Kay-Schuttleworth

Kay-Schuttleworth fue uno de los médicos activistas de la *Sociedad Estadística de Manchester*, que ocupó el cargo de secretario del Consejo de Sanidad Pública de Manchester. Este cargo le sirvió para recabar los datos necesarios para su investigación, sirviéndose de una red centralizada de inspectores sanitarios (que él mismo creó).

En esta investigación el autor describe las condiciones infrahumanas y de insalubridad en las que vivían las capas más desfavorecidas de la población inglesa de principios del siglo XIX.

c) Vida y trabajo de los habitantes de Londres (1889-1891), de Charles Booth

Una voluminosa obra, en tres volúmenes, publicados de 1889 a 1891, y que con posterioridad se editaría en diecisiete volúmenes. En ella se describe, con minuciosidad, las condiciones de vida de la población trabajadora londinense. La información manejada procede de diversas fuentes: los *censos de población,* realizados en Inglaterra hasta el momento de la investigación, y diversos informes (policiales, de organizaciones de beneficencia, de inspectores escolares y sanitarios).

En lo metodológico, este estudio destaca por tres aspectos importantes:

1) La *observación directa* de barrios obreros.
2) La realización de *entrevistas abiertas* a hombres de negocios.
3) La elaboración de *mapas sociales* de diferentes barrios de Londres. Mediante estos *mapas*, el autor analiza la distribución de la población en la ciudad. De

manera especial, la relación existente entre la ubicación de la industria y el nivel de vida de la población.

La figura de Booth destaca, no sólo por su obra, sino también por los seguidores que tuvo a comienzos del siglo XX. Entre ellos suelen citarse los siguientes:

a) *Rowntree.* A este autor se le reconoce un cierto perfeccionamiento de la *metodología de los sondeos:* preguntar directamente a los obreros, sin acudir a intermediarios.
b) *Bowley y Burnett.* Introdujeron mejoras en el *muestreo aleatorio.* Estas se centran en la disminución de los *errores muestrales* y de los gastos en su realización.
c) *El matrimonio Webb.* Principalmente por el análisis que efectuaron (mediante *fuentes secundarias)* de la evolución de determinadas instituciones sociales. De su obra debe mencionarse el texto clásico de Beatrice Webb *My Apprenticeship,* de 1926, y su obra conjunta *Methods of social study,* de 1936. En ella, los autores sintetizan su experiencia en la investigación social.

• De las *encuestas sociales* efectuadas en Francia merecen mencionarse cuatro fundamentales:

a) *Higiene social (1836)* y *La prostitución en la ciudad de París (1834),* de A. Parent-Duchatelet

En estas investigaciones se utilizaron, igualmente, diversas fuentes de información: *archivos policiales, datos estadísticos, entrevistas* y *observación personal.*

b) *Estadística sobre la moralidad en Francia (1833),* de G. Chanpneuf

En esta obra se analiza la relación existente entre las variables tasa de delincuencia y nivel cultural, entre otras variables.

c) *Obreros europeos (1877-1879),* de Frederic Le Play

Un ingeniero de minas interesado en el análisis de la familia como núcleo de toda la sociedad. Estudia no sólo las características externas de las familias, sino también sus sentimientos, valores y deseos. Para ello, compagina la *observación participante* con la *comparación sistemática* de diversas familias de países diferentes. [illegible]

le fue posible gracias a su profesión: responsable de la minería y metalurgia en varios países.

Nisbet (1966: 61) califica esta investigación como "una obra estrictamente sociológica, el primer trabajo sociológico genuinamente científico del siglo XIX". Sea o no cierta esta afirmación, parece indudable la contribución de Le Play a lo que después se denominaría el *estudio de casos.*

En España, la obra de Le Play fue elogiada por intelectuales de la talla del marqués de Pidal (quien dedica a Le Play su discurso de ingreso en la Real Academia de Ciencias Morales y Políticas, en 1887), R. Rodríguez de Cepeda, Jaime Balmes, Joaquín Costa, Azcárate y Posada. Estos tres últimos autores llevaron a cabo investigaciones sobre usos y costumbres de la población, con una intención claramente reformista.

• Aparte de las cuatro *encuestas sociales* (sucintamente comentadas), en Francia también adquieren una gran repercusión las realizadas por el jurista André-Michel Guerry: *Ensayo sobre la estadística moral de Francia (1832)* y *La estadística moral en Inglaterra comparada con la de Francia (1860).* Principalmente, por su contribución al desarrollo de las estadísticas de delincuencia.

Asimismo, sobresalen los estudios de Villermé. Un antiguo médico del ejército de Napoleón y máximo exponente de la tendencia *reformadora* en Francia. Este autor publicó varios estudios sobre problemas de higiene y sanidad. También realizó varias investigaciones acerca de las condiciones de vida en las cárceles francesas del siglo XIX.

• La gran evolución que experimenta la investigación social empírica a lo largo del siglo XIX fue pareja a los avances producidos en las técnicas de recogida y análisis de la información. Adolf Quételet (1796-1874) fue uno de los más importantes impulsores de la aplicación de las mejoras introducidas en el campo de la estadística al estudio de los fenómenos sociales. Colabora en la creación de las *sociedades estadísticas* de Inglaterra y Francia; e impulsa la fundación de la *Asociación Estadística Internacional*, para la cooperación internacional en la recogida y análisis de información social.

Pero, si por algo destaca Quételet es, sobre todo, por su *teoría de las regularidades de los fenómenos sociales:* en la conducta humana, aparentemente azarosa, pueden observarse regularidades, que pueden medirse mediante técnicas estadísticas de análisis. De ahí que se le considere "un pionero en la aplicación de modelos matemáticos en la sociología" (García Ferrando, 1979: 33).

En sus investigaciones socio-antropológicas *(Sobre la ley del crecimiento del hombre* e *Investigaciones sobre las tendencias criminales en distintas edades),* Quételet analiza las características físicas de las personas y las estadísticas de delincuencia. Para ello aplica el *cálculo de las probabilidades estadísticas.*

En 1835 publica su obra más relevante: *Sobre el hombre y el desarrollo de las facultades humanas: ensayo sobre física social.* Años más tarde (en 1869), reeditaría es-

ta obra, corregida y ampliada, pero con un título más breve: *Física social*. De ella se ha dicho que "marcó la transición de la simple descripción estadística al empleo consciente de los datos cuantitativos empíricos, para establecer las regularidades de la vida social" (Kovaliova, 1989: 136).

De sus indagaciones Quételet extrae el concepto de *hombre medio:* el término medio de las características que resultan de la observación de diferentes sujetos.

Además de Quételet hubo otros autores que realizaron importantes aportaciones en el campo de la *estadística*. Entre ellos destacan:

a) Laplace, que introduce el principio de mínimos cuadrados en su *Teoría analítica de las probabilidades*.
b) Gauss generaliza el *método de mínimos cuadrados*.
c) Yule aplica, por primera vez, los análisis de *regresión múltiple* y de *correlación* a datos sociales. Esto sucede a finales del siglo XIX.

En Alemania, los enfoques de Le Play y Quételet fueron sintetizados por tres autores principales:

a) Ernst Engel, en sus investigaciones sobre el presupuesto.
b) Adolf Wagner, considerado epígono de Quételet. A este respecto es muy ilustrativo el título de uno de sus libros: *La regularidad de las acciones que parecen casuales desde el punto de vista de la estadística*, de 1864.
c) Wilhelm Lexis, que profundiza en las ideas de Quételet sobre la cuantificación de los fenómenos sociales.

• La recogida y análisis de datos sociales poco a poco va profesionalizándose. El número de investigadores *amateurs* se reduce y aumenta, por el contrario, el de profesores universitarios. Entre éstos cabe citar a Gustav Schmoller, Ferdinand Tönnies y Max Weber, en Alemania; y, a Emile Durkheim, en Francia.

Esta profesionalización de la investigación social contribuyó a que el divorcio entre la *teoría* y la *empírica*, que había caracterizado a la investigación social empírica precedente, desapareciera. Si bien, no es hasta el siglo actual cuando se plantea, realmente, la necesidad de articular la *sociología empírica* con la *teórica*. A ello contribuyó Kuhn (1978: 132) cuando afirma, en sentido genérico, que "la ciencia normal se esfuerza y deberá esforzarse continuamente por hacer que la teoría y los hechos vayan más de acuerdo".

Cuando Kuhn habla de "ciencia normal", se está refiriendo a una "investigación basada firmemente en una o más realizaciones científicas pasadas, realizaciones que alguna comunidad científica particular reconoce, durante cierto tiempo, como fundamento para su práctica posterior" (Kuhn, 1978a: 33).

1.1.2. Principales contribuciones teórico-metodológicas de los fundadores de la sociología

En sus inicios, la sociología se caracteriza por ser una ciencia eminentemente teórica. En ella, la *empíria* aparece subordinada a la *teoría.* Se reconoce la importancia de la investigación, pero como un instrumento para reforzar la *teoría;* y no como un medio para su contraste o *verificación empírica.*

Rodríguez Ibáñez (1989) sintetiza la evolución de la teoría sociológica en dos líneas fundamentales:

a) *Estructural (dimensión gesellschaft),* que va del positivismo clásico, y el evolucionismo a la teoría de sistemas, pasando por Pareto y, sobre todo, por la reformulación sociológica de Durkheim y el funcionalismo.
b) *Intersubjetiva (dimensión gemeinschaft),* que arranca de movimientos no estrictamente sociológicos (como el romanticismo y la propuesta germana de *ciencias de la cultura).* Esta propuesta encuentra en Max Weber un refinador sociológico crucial (junto a otras figuras coetáneas, como Tönnies y Simmel); y, años después, en los representantes de la *teoría crítica* y de las corrientes fenomenológicas.

Los orígenes de la *dimensión estructural* los sitúa en la obra de Montesquieu, en su insistencia en las *leyes* naturales y sociales; así como en el equilibrio de los diversos órdenes. Rousseau será, en cambio, el promotor de la *dimensión intersubjetiva,* al precisar el fundamento individual o de libres voluntades en que descansa el *contrato social.*

De las dos dimensiones, la *estructural* es la predominante en los inicios de la sociología. Primero, bajo el nombre de *positivismo,* de la mano de Comte y de su maestro Saint-Simon. Después, con el nombre de *evolucionismo,* encabezado por Spencer.

• Comte toma de Montesquieu el determinismo de los fenómenos sociales e históricos; y de Condorcet, la afirmación de que el progreso del espíritu humano constituye la razón del devenir de la sociedad. Su objetivo principal será la reconstrucción del orden social, tras los cambios acaecidos con la venida de dos revoluciones: la francesa y la industrial. Ello demandaba el nacimiento de una nueva ciencia. Con la Revolución francesa, las concepciones metafísicas y teológicas del mundo habían perdido vigencia. El nuevo orden instaurado exigía, por el contrario, una ciencia positiva de la sociedad.

Al principio Comte denomina a la nueva ciencia *Física social,* por su equiparación a las ciencias de la naturaleza. Pero, años después, en 1839, en el IV volumen de su obra *Curso de filosofía positiva,* aparece con la que sería su denominación definitiva: *sociología.* El cambio de nombre responde al deseo de Comte de diferenciar a la nueva ciencia de la *física social* descriptiva de Quételet (quien por aquel entonces acababa de publicar su obra *Ensayo de física social).* Además, la intención de Comte era hacer de la nueva ciencia una disciplina teórica. Tendría por objeto el descubrimiento de

las leyes fundamentales inherentes a los fenómenos sociales. Para ello se basaría en dos *postulados metodológicos* esenciales:

a) Observar los fenómenos sociales desde un punto de vista científico, a semejanza de los fenómenos naturales.
b) Analizarlos, no como hechos aislados e independientes unos de otros, sino como fenómenos entrelazados y comprendidos en un todo.

En su *Curso de filosofía positiva,* Comte afirma que la *observación* constituye el método principal de la investigación social. A éste añade el *experimento,* el *método comparativo* y el *histórico*. Este último consistiría en la comparación histórica de los distintos estadios consecutivos de la humanidad. Pero, por desgracia, sus reflexiones metodológicas no quedaron plasmadas en investigaciones empíricas concretas. Todo quedó en una extraordinaria especulación, aunque de gran repercusión en el desarrollo ulterior de la sociología.

• El positivismo de Comte da paso a un período de la sociología, en el que la visión estructural-funcional de la sociedad se conjuga con la visión histórico-evolucionista. Spencer (1820-1903) fue, sin duda, el máximo exponente de esta etapa. Su tesis fundamental es que todo en el mundo se rige por una ley de la evolución, tanto la materia orgánica como la propia sociedad. Por tanto, son dos las ideas principales que estructuran su pensamiento:

a) La comprensión de la sociedad como un *organismo* (asemeja los organismos sociales a los biológicos).
b) La idea de la *evolución social.*

Con Comte comparte las nociones siguientes:

a) La *unidad de la ciencia.* Spencer también concibe la *sociología* como una *ciencia positiva,* basada en los mismos principios metodológicos de las ciencias naturales.
b) El concepto de *ley científica:* todo fenómeno está sujeto a regularidades, y su identificación se convierte en el objetivo de cualquier análisis.
c) El *holismo:* la comprensión de los hechos o fenómenos sociales como partes de un todo.
d) La restricción del conocimiento a la *experiencia,* a lo observable.

A Spencer se debe el primer manual sobre el *concepto* y el *método* de la *sociología*. Se trata de su obra *El estudio de la sociología* (1873) que, posteriormente, constituiría el prefacio de su voluminoso libro *Principios de sociología* (1877). Esta última

publicación se considera "el primer intento sistemático de desarrollar el análisis sociológico" (Lucas Marín, 1982).

• Al igual que Comte y Spencer, Pareto (1848-1923) propugna, en su *Tratado de sociología general* de 1916, la necesidad de reconstruir el todo social para alcanzar una explicación sintética del conjunto de la sociedad y de los movimientos que en ella se manifiestan. Nada de lo que sobrepasa el ámbito de la experiencia puede ocupar un lugar en la ciencia. La sociología se constituye, de este modo, en una ciencia empírica, inductiva y lógico-experimental, en la que debe imperar la neutralidad valorativa (al menos desde un punto de vista metodológico). Pero, aunque Pareto propugna la necesidad de realizar investigación empírica, sus reflexiones, al igual que las de sus predecesores, no aparecen acompañadas de datos empíricos que las corroboren.

• Por el contrario, Tönnies (1855-1935) sí lleva a cabo distintos estudios empíricos sobre conducta desviada y suicidio, aunque no logra su conjugación con su teoría de la sociedad. Hecho que sí se alcanza con Durkheim (1855-1917). Con este último autor culmina la visión orgánica de la sociedad, junto con la consideración positiva de la sociología. Se convierte en el primer representante de la *perspectiva estructural-funcionalista,* y el primer autor que define el procedimiento que debe seguirse para alcanzar el conocimiento científico de los fenómenos sociales.

En 1895 Durkheim publica *Las reglas del método sociológico*. De este clásico texto se ha dicho que "no son propiamente unas reglas del método –unas *técnicas*, como se diría luego–, sino unos verdaderos principios de la epistemología sociológica" (Marsal, 1977: 146). En esta obra se articula su visión del *método* y del *objeto* de la *sociología*. Quizás la propuesta más característica que se halla en esta obra sea la concepción de la sociedad como una suma de *hechos,* que han de analizarse como *cosas*. Ello supone el tratarlos como *data,* el analizarlos desde fuera, como cosas exteriores, y descartar todas las nociones previas.

En opinión de Durkheim, el estudio de la realidad social ha de incluir dos *tipos de análisis:* el *causal* y el *funcional*. Si se quiere explicar un *hecho social* hay primero que buscar su *causa*, definir el fenómeno social que le antecede para, una vez establecida la *causa*, indagar la *función* que cumple en el sistema social (su utilidad). El *método* idóneo es el *comparativo,* "el único que conviene a la sociología" (Durkheim, 1974: 133). Sólo podremos descubrir lo que es específico de una sociedad y lo que es común a la humanidad, si comparamos los mismos fenómenos en diferentes sociedades.

Pero, sin duda, el gran mérito de Durkheim lo constituye el haber aplicado de forma coherente las *reglas del método* que propone a investigaciones concretas: *El suicidio* (1897) y *Las formas elementales de la vida religiosa* (1912). En concreto, de su obra *El suicidio,* Caplow (1977: 10) afirma que

> "Ha sido la primera obra importante que ha configurado la tradición de la teoría social con la aritmética política (...). Fue el prototipo perfecto de informe de investigación moderna (...). La parte teórica está bastante desarrollada, ya que el interés teórico del

autor supera el problema del suicidio en sí. Las teorías anteriores son examinadas detalladamente, y el marco de la investigación está claramente definido; el significado de las estadísticas de suicidios es estudiado en todas sus ramificaciones; las hipótesis son modificadas y mejoradas como consecuencia de las confrontaciones con los datos, y reciben una nueva formulación al final de la obra."

Entre los seguidores más notorios de Durkheim destacan los antropólogos funcionalistas Malinowski y Raddcliffe-Brown. Estos antropólogos aplicaron las hipótesis durkheimianas (de la estructura normativa unitaria de la sociedad y su variación interna de funciones), a sus respectivas investigaciones sobre comunidades primitivas. Igualmente destacan sociólogos y psicólogos sociales como Parsons, Merton, Festinger, Duncan y un largo etcétera.

Pero no todos fueron seguidores. Las tesis de Durkheim también encuentran oponentes. Entre ellos destaca muy especialmente Max Weber (1864-1920) y su *sociología comprensiva*.

• A diferencia de Durkheim, Weber parte de la distinción propuesta por Dilthey entre *ciencias naturales* y *ciencias del espíritu*. Las primeras se centran en la observación de las manifestaciones externas de los fenómenos naturales, en tanto que las segundas, en la interpretación del *significado* y el *sentido* de la *acción* humana. Esta interpretación es posible gracias a un proceso de *comprensión (verstehen):* el ponerse en el lugar del otro.

Weber se adhiere a la *escuela neokantiana*. Reconoce que se puede *comprender* una buena parte de las *acciones humanas* por *medios empáticos* (compenetración, intuición). Pero, prefiere considerar la *acción humana* como racional, orientada a *valores;* e interpretar el *sentido subjetivo* inherente a la misma.

La tesis de Rickert sobre los *conceptos* como medios para superar la diversidad de la realidad empírica se transforma, en la teoría de Weber, en la noción de *tipo ideal*.

Weber define el *tipo ideal* como el "interés de la época", expresado en forma de construcción teórica. No se extrae de la realidad empírica, sino que se construye como un esquema teórico "mediante el realce unilateral de uno o de varios puntos de vista y la reunión de una multitud de fenómenos singulares, difusos y discretos, que se presentan en mayor medida en unas partes que en otras o que aparecen de manera esporádica fenómenos que encajan en aquellos puntos de vista, escogidos unilateralmente, en un cuadro conceptual en sí unitario" (Weber, 1973: 79).

Weber considera que la "construcción de tipos ideales abstractos no interesa como fin, sino exclusivamente como medio" (Weber, 1977: 63). Puede servir como patrón de referencia para comparar con él la realidad empírica. Si bien, puntualiza que el fin de la formación de conceptos de *tipo ideal* consiste no en tomar conciencia de lo genérico, sino de la particularidad de los fenómenos sociales. Mediante los *tipos ideales* se puede comprender la *motivación* de los actores. También puede establecerse conexiones entre la *motivación* y los resultados de la *acción*.

Weber se centra en el aspecto *cualitativo* de los *hechos sociales*. Cuestiona la aplicación que del concepto de *ley* se ha hecho, tanto en las ciencias naturales como en las sociales. En su opinión, el establecimiento de *leyes* no constituye el objeto de la ciencia, sino que representa una fase previa al conocimiento que aspiramos. Cuatro serían en total, los *pasos* a seguir en la adquisición del conocimiento:

1) Fijar conceptos y establecer regularidades.
2) Agrupar causas posibles en cada caso y momento histórico determinado, señalando su importancia y a qué se debe.
3) Remontarnos al pasado para ofrecer una explicación histórica de las distintas particularidades de tales agrupaciones.
4) Evaluación de las posibles constelaciones de factores en el futuro (Weber, 1977: 41).

Asimismo, Weber no acepta el principio de la *causalidad:* "el conocimiento de unas leyes de la causalidad no puede constituir el fin, sino el medio de estudio" (Weber, 1977: 45). Aunque sí defiende, también, la necesidad de que la sociología sea una ciencia objetiva, neutral y libre de valores. El científico se diferencia del político en que debe explicar *objetivamente* los hechos, sin implicaciones normativas o valorativas.

Pero, a diferencia de Durkheim, Weber no desarrolla su esquema metodológico general en una obra concreta. Su primer trabajo empírico data de 1892, y lo titula *La situación de los obreros de la industria textil en Alemania al este del Elba.* Se trata de una investigación empírica que Weber realiza en una fábrica textil alemana. Años más tarde, publica una introducción metodológica sobre aquella investigación *(Sobre el problema de los fundamentos psicofísicos del trabajo industrial)*. Sus intereses metodológicos coincidieron en dos temas fundamentales: "la elección de los informantes y la mejora del diseño del cuestionario" (García Ferrando, 1979: 62).

La obra de Weber fue determinante en el desarrollo de la *sociología cualitativa*. En concreto en:

a) El *interaccionismo simbólico* posterior a Mead (Blumer, Glaser y Strauss).
b) La investigación *fenomenológica* de Schütz, Husserl, Berger y Luckman.
c) La *etnometodología* o sociología cognoscitiva de Garfinkel, Douglas, Cicourel y Lindzey.

En el Cuadro 1.1 se resumen algunas de las características definitorias de estas dos formas principales de analizar la realidad social.

• A caballo entre estas dos dimensiones se encuentra la obra de Karl Marx (1818-1883) y su método del *materialismo dialéctico*. Consiste en fijar las *leyes* que regulan

CUADRO 1.1. Características fundamentales de dos dimensiones básicas en el análisis de la realidad social.

Dimensión estructural	*Dimensión intersubjetiva*
Analogía metodológica ciencias sociales con ciencias naturales: unidad de la ciencia.	Diferenciación ciencias sociales de las ciencias naturales: singularidad de su objeto y método de conocimiento.
Búsqueda de *regularidades* objetivas de los fenómenos sociales: identificación de *leyes universales*.	Análisis de lo *individual* y *concreto*. Énfasis en la variedad y especificidad de las sociedades humanas y de sus manifestaciones culturales.
Descripción y *explicación* de los *hechos sociales* como elementos integrantes de la sociedad global. Interesan los *aspectos cuantitativos*.	*Comprensión* o *interpretación* de los agentes internos de la *acción social*. Destacan los *aspectos cualitativos*.
Énfasis en la *cohesión estructural*.	Se centra en la *actividad intersubjetiva*, concediéndo gran importancia a la *historia* (a la condicionalidad histórica y su significado cultural).
Positivismo, evolucionismo, funcionalismo, teoría de sistemas...	Romanticismo, historicismo, escuela neokantiana, fenomenología, interaccionismo simbólico, etnometodología...

los fenómenos sociales y, al mismo tiempo, enfatizar la perspectiva *histórica*. Así, por ejemplo, en su voluminosa e incompleta obra *El capital*, Marx utiliza como fundamento de sus argumentaciones tanto datos estadísticos como documentos históricos.

La influencia de Marx en el desarrollo de la teoría y de la investigación empírica sociológica también ha sido notoria e incesante. Por un lado, su obra ha inspirado diversos estudios: desde los clásicos de Lazarsfeld, Berelson *et al.* (1944) sobre la influencia de la clase media en el voto, o los de Lockwood (1958) y Goldthorpe (1980) acerca del aburguesamiento de la clase obrera; hasta las líneas de investigación más recientes acerca de la estratificación social y otras tantas investigaciones derivadas, directa o indirectamente, de la obra de Marx.

Por otra parte, no debe olvidarse la influencia de Marx en sociólogos posteriores; aunque, en unos casos fuese para revisarlo y, en otros, para criticarlo. A este respecto, puede citarse a Luckàcs o a Gramsci, entre los autores que han desarrollado aspectos inacabados de la obra marxiana; o a Althusser y Poulantzas, como representantes de la versión estructuralista del marxismo; y a los miembros de la *teoría crítica* de la Escuela de Frankfurt (Horkheimer, Marcuse, Adorno...), entre otros.

1.2. Recepción, institucionalización y crisis en el desarrollo de la sociología: implicaciones metodológicas

Hasta los años veinte del presente siglo los desarrollos teóricos primaron sobre las aportaciones empíricas. A partir de entonces es cuando se plantea, realmente, el problema de la fusión de la *sociología teórica* con la *empírica*. Esto coincide con los inicios de la institucionalización académica de la sociología. En ella Estados Unidos adquiere el protagonismo, practicando una *sociología* más *empírica* que *teórica*.

A comienzos del siglo XX, la situación en Europa era muy dispar. En Inglaterra, el centro de la sociología lo representaba la *Sociedad de Londres* (fundada en 1903) con la divulgación, desde 1908, de la revista científica *Sociological Review*. En 1907 se crea la primera cátedra de sociología en la Universidad de Londres, en la figura de Hobhouse.

En Francia la institucionalización de la sociología como asignatura académica acontece en 1896, cuando Durkheim ocupa la primera cátedra de "ciencia social" en la Universidad de Burdeos. Sin embargo, en 1917, con el fallecimiento de Durkheim desaparece también *L'Année Sociologique,* la única escuela que podría haber proseguido los trabajos empíricos de Durkheim.

En Alemania la situación era peor. Pese a que en 1909 se crea la *Sociedad Sociológica Alemana,* encabezada por Tönnies, la animadversión que la facultad de filosofía (donde se hallaban concentradas todas las ciencias humanistas) sentía hacia las *indagaciones empíricas,* contribuyó a que la institucionalización de la sociología en Alemania fuese tardía, respecto a otros países.

Tras el fallecimiento de Weber en 1920, el Círculo de Viena y la Escuela de Frankfurt centran la teoría sociológica alemana. Pero, sólo hasta la llegada del nazismo, que supuso su paralización y expulsión a otros países. Estados Unidos será el país de acogida fundamental, convirtiéndose en el centro de la sociología mundial.

Por último, en España la institucionalización de la sociología fue anterior a otros países europeos. En 1898, se crea la primera cátedra de sociología en la antigua Universidad Central (hoy Universidad Complutense), dentro de la Facultad de Filosofía y Letras. La plaza la ocupará Manuel Sales y Ferré. Pero, a su pronta institucionalización no le sigue una rápida maduración. Esta no dará comienzo hasta los años cincuenta, y bajo la influencia principal de la sociología norteamericana. Hasta entonces, habían destacado las aportaciones de cuatro teóricos de la sociología:

a) Gumersindo de Azcárate (1840-1917), que introduce el *evolucionismo* de Spencer en España y funda el *Instituto Internacional de Sociología* en 1891.
b) Manuel Sales y Ferré (1843-1910), que edita el primer manual de sociología en España, el *Tratado de sociología. Evolución social y política* (en cuatro volúmenes, de 1889 a 1897), y el primer tratado de *Sociología general* (1912), dentro de la corriente evolucionista.
c) Adolfo G. Posada (1860-1944) y sus *Principios de sociología.*
d) Ortega y Gasset (1883-1955), alrededor de quien se constituye la *Escuela de Madrid.*

La Guerra Civil eclipsa estos primeros balbuceos de la sociología en España. Hasta que, en los años cincuenta, Gómez Arboleya comienza su refundación.

• Si Europa fue la cuna de la sociología, Estados Unidos supone su cristalización como ciencia empírica. A ello contribuyó, de acuerdo con Shils (1971), el asociacionismo voluntario americano (con su capacidad económica para financiar programas de investigación social empírica), junto a la creación de nuevas universidades. En ellas el establecimiento de departamentos de sociología no encuentra la fuerte oposición de las disciplinas "clásicas" (la "vieja guardia"), como sucedía en las universidades europeas, caso de las alemanas. A estos factores se sumaron otros, entre los cuales Kon (1989) destaca los siguientes:

a) La influencia del pragmatismo y el spencerismo.
b) La competencia de las universidades.
c) El amplio movimiento a favor de las reformas sociales.

En los subapartados siguientes se detallan el caso estadounidense y el español.

1.2.1. El caso estadounidense

En Estados Unidos, el primer curso que lleva el nombre de *sociología* lo dicta Sumner en la Universidad de Yale en 1876. Aunque la publicación de los dos primeros libros, donde aparece la palabra *sociología* en su título, data de 1854. Se trata de *Sociology of south,* de Fitzbuth, y *Treatise on sociology,* de Hughes.

En 1893 Small ocupa la primera cátedra de sociología en la Universidad de Chicago. En 1895 aparece el primer número de la revista *American Journal of Sociology*. Y, en 1905 se constituye la *Sociedad Sociológica Norteamericana*, presidida por Ward, fruto de su escisión de la *Sociedad Económica Norteamericana*. A continuación, la creación de departamentos de sociología se suceden en diferentes universidades: Columbia, Wisconsin, Minessota, Míchigan. Con ello se acentúa la especialización de la sociología, con un marcado carácter científico.

• Esta nueva etapa de la sociología se estrena con investigaciones ambiciosas, en las cuales se trata de conjugar la *teoría* con la *empíria*. Entre esas investigaciones destaca *El campesino polaco en Europa y América (1918-1920),* de W. Thomas y F. Znaniecki.

Esta obra supuso la introducción de nuevas técnicas de recogida y de análisis de datos, no habituales en la investigación empírica tradicional. Concretamente, la recopilación y análisis de *documentos personales*. El objetivo del estudio era observar el proceso de adaptación social de los inmigrantes polacos en Estados Unidos, a principios del siglo XX. Para ello, los autores se basaron en una gran diversidad de fuentes de datos (754 cartas de inmigrantes polacos en Estados Unidos; archivos de un periódico ad-

quirido por Thomas en Polonia, y que cubría un período de 20 años; archivos de parroquias polacas en Estados Unidos; documentos de algunas organizaciones de ayuda a inmigrantes polacos; y la autobiografía de un joven inmigrante polaco llamado Wladeck). Todo lo cual convierte al estudio en original y pionero en la aplicación del *método biográfico*.

Siguiendo la línea marcada por Thomas y Znaniecki, aunque desde una perspectiva ecológica, se encuentran los trabajos de la *Escuela de Chicago,* dentro del marco del Departamento de Antropología y Sociología de dicha Universidad. Park –que por aquel entonces dirigía ese departamento universitario– realiza y dirige (bajo la influencia de Simmel, a cuyas clases asistió en Berlín) una serie de investigaciones centradas en la influencia del entorno (generalmente urbano) en la organización e interacción social. Propone el estudio del barrio, su composición, organización y relación con el resto de la comunidad.

En 1916 Park publica (para la revista *The American Journal of Sociology)* un artículo titulado "La ciudad: proposiciones para el estudio del comportamiento humano en el medio urbano". A partir de este artículo se suceden numerosas investigaciones, publicadas en las *Chicago Sociological Series* o en la revista *The American Journal of Sociology.*

Esta última revista se convierte en el órgano de divulgación de los estudios de la Universidad de Chicago. Entre esos estudios sobresalen los siguientes: *The Gang,* de Tharsher; *The Jack Roller* y *Brothers in Crime,* de Shaw; y *The Ghetto,* de Louis Wirth.

La *metodología* seguida en éstos y en otros estudios es similar. A la *observación participante*, considerada como la técnica principal, añaden toda una variedad de datos sociológicos procedentes de *diversas fuentes: informes oficiales* (municipales, de los tribunales de justicia, de la policía, de las asociaciones locales), *datos censales, documentos privados y biografías, artículos de prensa, informes de especialistas, entrevistas y encuestas,* fundamentalmente.

• Entre las investigaciones que complementan la *observación participante* con otras *técnicas cualitativas* y *cuantitativas* destaca, también, la realizada por el matrimonio Lynd: *Middletown* (1929 y 1937). Una investigación calificada como "el primer *best-seller* sociológico" (Caplow, 1977: 39).

Durante dieciocho años, Helen y Robert Lynd recogieron datos estadísticos y participaron, en la medida de sus posibilidades, en la vida de una pequeña ciudad industrial de 40.000 habitantes (de Indiana), llamada, en realidad, Muncie. El objetivo del estudio era comparar el *Middletown* contemporáneo (1924-1925) con el de hacía una generación. Para ello utilizaron toda una variedad de *documentos:* archivos de la ciudad y del condado; informes de los tribunales de justicia; memorias anuales de la comunidad y de asociaciones locales; diarios locales e, incluso, diarios personales de dos hombres que se ofrecieron voluntarios al estudio.

Además, realizaron *entrevistas abiertas* y *estructuradas,* mediante un *cuestionario,* a distintos habitantes seleccionados de la ciudad. Todo lo cual contribuyó a que este estudio *(Middletown),* de 1929, y su continuación *(Middletown in Transition),* de

1937, supusieran un gran impulso al empleo de la *entrevista informal* en investigaciones posteriores sobre comunidades.

• La serie *Yankee City (1941-1959)*, de W. Lloyd Warner, sería una de estas investigaciones. Se desarrolló durante cinco años (de 1931 a 1935), en la ciudad de Newburyport (Massachusetts). Una pequeña ciudad costera de 17.000 habitantes, cuya vida social el autor describe en los cinco gruesos volúmenes que componen la obra.

Este estudio guarda bastante similitud con su antecesor, *Middletown*. También, en este caso, el autor se establece en la ciudad y participa en la vida de la comunidad. Igualmente, reúne una gran cantidad de información de diferentes fuentes: observación participante, documentos oficiales, periódicos, diarios personales y estudios de casos particulares.

Las investigaciones de *Yankee City* derivaron, asimismo, de las experiencias de Hawthorne, en las que Warner participó, aunque de manera esporádica. Dichas experiencias, conocidas también con el nombre de investigaciones en la *Western Electric*, fueron realizadas desde 1927 hasta 1932, bajo la dirección conjunta de dirigentes de la Western Electric Company y un equipo de investigadores de Harvard, dirigidos por Elton Mayo.

El informe definitivo de la investigación lo publicaron Roethlisberger y Dickson en 1939 *(Management and the Worker)*. Pero algunos años antes, Mayo había incluido un resumen de los resultados en su obra *The Human Problems of an Industrial Civilization*. Esta última publicación destacó, principalmente, por ser la primera en la que se hace un *análisis de contenido* de amplias dimensiones (a partir de 80.000 proposiciones extraídas de las entrevistas realizadas durante 1929); y por la utilización de *entrevistas no dirigidas* (o *no estructuradas)*.

• La obra de William Foot Whyte, *Street Corner Society* (1943), también tuvo una gran repercusión metodológica. Esta obra significó la culminación de la técnica de *observación participante* en la investigación social. A diferencia de las investigaciones anteriores, el autor fue sólo ayudado por su mujer y un condiscípulo, y únicamente utilizó la *observación participante* como técnica de recogida de información fundamental.

En 1955 se publica el *Apéndice metodológico* de la obra. En él, Whyte narra los pasos seguidos hasta su aceptación en Corneville (un barrio marginal de Boston, poblado por italianos), así como las incidencias del estudio.

• Con la llegada a EEUU, a finales de los años treinta y principios de los cuarenta, de algunos miembros del *Círculo de Viena* (que huían de la persecución nazi), la sociología dominante hasta entonces cambia de rumbo. La confluencia del *positivismo lógico* del *Circulo de Viena* con el *pragmatismo americano,* junto a las críticas manifestadas contra la *metodología cualitativa*, supusieron el fin del dominio de la *Escuela de Chicago*, y el auge de la *metodología cuantitativa*. La Universidad de Columbia tomó el relevo a la de Chicago.

En Columbia estaban los dos sociólogos principales del momento: Robert K. Merton y Paul F. Lazarsfeld. Con ellos "la tradición teórica y empírica se aproximaron más que en otro lugar, pero sin llegar a producirse, tampoco, una integración total" (Shils, 1971: 159).

El desarrollo de la *metodología cuantitativa* fue posible gracias a las aportaciones de matemáticos, como el londinense K. Pearson *(The Grammar of Science,* 1892), o el francés Poincaré (1854-1912); y, en especial, del instrumentalista John Dewey (1859-1952). Este último influyó en la obra de Lundberg y en el razonamiento teórico metodológico de la *sociología positiva* de la época.

Asimismo, García Ferrando (1979) cita, también como antecedente, el artículo de L. L. Thurstone *"Attitudes Can Be Measured"* (1928). Este artículo "representa un punto de inflexión significativo en el desarrollo de la sociología empírica contemporánea" (García Ferrando, 1979: 79).

De las investigaciones realizadas bajo la influencia del *operacionalismo* y del *positivismo instrumental* resaltan dos fundamentales:

a) *The People's Choice: How the Voter Makes Up His Mind in a Presidential Compaing* (1944), de P. Lazarsfeld y Berelson

Esta obra destaca, sobre todo, porque introduce los *estudios panel* en la investigación social.

El objetivo de la investigación era analizar las intenciones de voto de los electores del condado de Erie (Ohio), en las elecciones presidenciales de 1940. Para ello, los autores pasaron tres veces un *cuestionario* a una *muestra* de electores extraída al azar. La primera vez, antes del nombramiento de los electores; la última, momentos antes de la elección.

A su vez, utilizaron otras tres *muestras* como *grupos de control,* con el propósito de comprobar los efectos de aplicar varias veces el mismo *cuestionario* a la *muestra* principal.

A partir de este estudio, las *encuestas* sobre intención de voto se ponen de moda.

b) *The American Soldier* (1949), de A. Stouffer *et al.*

Esta investigación destaca, en cambio, por la aplicación que en ella se hace de *datos secundarios*; además de las mejoras que introduce en el diseño y análisis de la *encuesta.* Concretamente, en la elaboración de varias preguntas para la medición de un mismo *concepto teórico;* e innovaciones en los *cuestionarios autoadministrados.*

• Los años cincuenta y sesenta se caracterizan por "la proliferación de fundaciones y de institutos, por la extensión de las subvenciones y del personal de investigación, por la creación de archivos estadísticos, por la introducción de ordenadores y por la generalización de las investigaciones sociológicas por todo el mundo. La consecuencia de todo ello fue un aumento de la especialización" (Caplow, 1977: 98).

A principios de los sesenta, el *funcionalismo* pierde la hegemonía entre los jóvenes sociólogos. Surgen nuevas perspectivas de la realidad social y renacen viejas teorías (algunas de ellas con nombres nuevos). Si bien, ninguna perspectiva llega a dominar el desarrollo de la sociología del momento.

Los importantes avances que la *metodología cuantitativa* experimenta en aquellos años (en gran medida propiciados por la aparición de los ordenadores, que permiten realizar análisis estadísticos hasta entonces imposibles), despiertan las viejas críticas de los sociólogos *cualitativistas*. Estos defienden la especificidad de las ciencias sociales, y la necesidad de utilizar técnicas de análisis diferentes a las aplicadas en las ciencias naturales.

Las antiguas disputas entre *positivistas* e *historicistas* resurgen de nuevo. Sirva de ilustración la pronunciación hecha por Adorno, al respecto, en 1969:

> "Hace mucho tiempo que la sociología empírica se ha dado cuenta de lo que pierde en contenido específico por culpa de la generalización estadística. A menudo viene contenido en un detalle algo decisivo desde el punto de vista de lo general y que se escapa a la mera generalización. De ahí la fundamental necesidad de contemplar los sondeos estadísticos por medio de *case studies*. El objetivo de los métodos sociológicos cuantitativos debería ser, asimismo, la comprensión cualitativa; la cuantificación no es un fin en sí misma, sino un medio para dicho fin" (Adorno 1973: 51-52).

Adorno subordina la *cuantificación* a la *comprensión cualitativa*. Este autor se encuadra, junto a Horkheimer, Marcuse, Bloch, o Habermas, dentro de la *teoría crítica*. En esta corriente convergen el pensamiento de Marx, con el de Weber y Freud. Se caracteriza por la crítica que realiza a la exaltación de la experiencia y de las ciencias empírico-analíticas, junto con la defensa de la *autorreflexividad*. Sus teóricos defienden la necesidad de que la ciencia social transcienda la *apariencia* de los fenómenos sociales, para ir hasta la *esencia* de las cosas.

Los orígenes de esta corriente crítica se remontan, no obstante, a los años treinta, cuando se constituye el *Instituto de investigación social de Frankfurt,* en torno a las figuras de Horkheimer (1893-1973), Adorno (1903-1969) y Marcuse (1897-1968), principalmente. De 1937 data la obra de Horkheimer *Teoría tradicional y teoría crítica*. Un manifiesto de la *Teoría crítica,* en el cual el autor define qué debe entenderse por *teoría crítica*.

En los años sesenta Habermas recoge el legado de los representantes de la Escuela de Frankfurt. Ello supone el renacimiento de sus planteamientos, pero en el marco de la crítica al *estructural funcionalismo*. En *Theory and Practice* (1974), Habermas sostiene que el *paradigma* ya no es la *observación cuantitativa,* sino el *diálogo*. En *disciplinas hemeneúticas,* como la *sociología,* ha de recurrirse al *diálogo,* y no a *procedimientos de medición*.

• En los años sesenta también se recupera la perspectiva del *Interaccionismo simbólico de la Escuela de Chicago* (Mead, Cooley, Thomas, Park), de la mano de algu-

nos de sus antiguos miembros. Entre ellos destacan Blumer, Becker, Hughes, Hanowitz y Goffman. Para estos autores, la sociedad es *interacción* y, como afirma Blumer (1966: 542), "el estudio de la interacción habrá de hacerse desde el punto de vista del actor. Dado que la acción la construye el actor a partir de lo que percibe, interpreta o juzga...."

Una reseña de las investigaciones realizadas desde el *Interaccionismo Simbólico* ha de incluir las aportaciones de Turner (1962) a: los problemas de identidad, el proceso de *role-making* y los movimientos sociales colectivos. Igualmente, los estudios de Rose (1962) sobre el poder y la misma teoría del interaccionismo simbólico; los análisis del comportamiento político simbólico de Edelman (1964); las investigaciones sobre la identidad y el contexto social de la muerte, de Strauss (1959); o, dentro de la teoría de la desviación, las investigaciones sobre el proceso de estigmatización y el desarrollo de la *carrera delictiva,* de los llamados teóricos del *etiquetamiento:* Becker (1963), Lemert (1969), Erikson (1962), Kitsuse (1962), o Schur (1968), entre otros.

• Otros sociólogos se orientarán hacia versiones de la *fenomenología,* como la *etnometodología.* El término *fenomenología* aparece, por vez primera, en la obra de Hegel, *Phenomenology of Mind.* Pero fue Edmund Husserl (1859-1938) quien configura la nueva corriente de pensamiento.

Este autor se sitúa dentro del *subjetivismo filosófico* de Descartes y Kant. Considera que es la conciencia la que crea el mundo, y no la experiencia del mundo la que produce el conocimiento. En 1913 publica su obra más importante *Ideas para una fenomenología pura y filosofía fenomenológica.*

Alfred Schutz (1899-1959) será en quién se basen las corrientes posteriores. Defiende que el mundo está compuesto de *significados.* El científico social deberá estudiar los "contextos significativos objetivos de contextos significativos subjetivos" (Schutz, 1972: 91).

Sin embargo, la expresión más clara de la *fenomenología sociológica* se encuentra en la obra de Berger y Luckmann (1968), *La construcción social de la realidad.* En ella los autores exponen cómo se construye el conocimiento y se orienta a algún problema social práctico. Los hechos que acontecen nunca son imparciales. Por el contrario, siempre reflejan una realidad.

En los años sesenta, en la Universidad de los Ángeles, surge, de la mano de Garfinkel, una nueva corriente de pensamiento denominada *etnometodología.* Su objetivo principal será traducir a categorías sociológicas operativas el programa filosófico de la *fenomenología.* El estudio del *lenguaje* (el desciframiento de los significados culturales) adquiere una gran importancia.

Los *etnometodólogos* analizan cómo las personas definen y construyen, cara a cara (en interacción), cada situación social. No rechazan el empirismo. Pero, proponen una nueva técnica: la *simulación.* Esta técnica consiste en partir de una situación que se cree normal y, desde ella, intentar crear confusión. Por lo que su operatividad en contextos que no sean microsociológicos se estima dudosa.

La *etnometodología*, en cambio, ha resultado fructífera en ámbitos concretos como la *sociolingüística (análisis del discurso)* y la *psicología social* de instituciones (juzgados, hospitales, escuelas, comisarías), en la ilustración de la realidad cotidiana de su funcionamiento.

Pero éstos no son los únicos desarrollos teóricos de los años sesenta. Junto a ellos conviven otros, como la *teoría de sistemas* (una actualización del *paradigma normativo),* o la *teoría del intercambio* (aunque los orígenes de esta última se remontan a 1907, cuando Georg Simmel publica un artículo que lleva por título el término "intercambio").

Los años sesenta se definen como los años de la heterodoxia, encarnada por disidentes (más o menos radicales), del *paradigma* que hasta entonces había imperado: el *cuantitativo*. A partir de los años setenta, comienza a hablarse de *crisis* de la sociología. *Crisis* que, favorablemente, trae un período de pluralidad paradigmática en la sociología.

1.2.2. El caso español

En el caso de la sociología española, la palabra clave es *recepción*. Sus etapas de desarrollo (surgimiento, resurgimiento, institucionalización y especialización) van a estar marcadas por lo que acontezca en el exterior y consiga importarse, dentro de unas fronteras nacionales que encierran un contexto histórico particular.

Al exilio, tras la Guerra Civil (en los años cuarenta), de buena parte de los padres fundadores/receptores, le sigue, en los años cincuenta y decenios posteriores, una segunda emigración científica: de las generaciones nacidas tras la Guerra Civil, que viajan a las universidades extranjeras (ante la penuria de la enseñanza de la sociología existente en España en aquel entonces). Estos se convierten, tras su formación o perfeccionamiento, en el vehículo humano de la *recepción*. Es el caso de dos de los discípulos de Gómez de Arboleya: Salustiano del Campo (que introduce en España el *funcionalismo* y el *empirismo norteamericano)* y Salvador Giner –que incorpora la *teoría social crítica*– (Lamo de Espinosa, 1992).

En España, en los años sesenta –al igual que en otros países– la sociología aparece bifurcada en dos grandes corrientes teóricas. Estas se configuran –siguiendo a Lamo (1992)– en una *vertiente empírica liberal,* encabezada principalmente por Juan Linz (Amando de Miguel, Juan J. Toharia, José Castillo, José Cazorla) y Francisco Murillo (Miguel Beltrán, José Jiménez Blanco y Juan Díez Nicolás); y el denominado *"pensamiento crítico"* de Luis González Seara, Juan Marsal y Carlos Moya (Victor Pérez Díaz, José M.ª Maravall, E. Lamo de Espinosa y Jesús Ibáñez).

Hasta entonces, la orientación de la sociología española había sido, a diferencia de la sociología americana, *teórica* más que *empirista*. Comienzan a crearse los primeros institutos de investigación de mercado y de sociología aplicada, cuyo desarrollo se acelera en los años ochenta.

El período comprendido entre 1963 y 1973 supone "la definitiva institucionalización de la sociología" en España (Lamo, 1992). En 1963 se crea el *Instituto de la Opinión Pública* (que en 1976 se transformará en el actual *Centro de Investigaciones Sociológicas).* Un centro oficial de investigación sociológica empírica, que encuentra en la publicación de la *Revista Española de la Opinión Pública* un medio de difusión.

En estos mismos años, destaca el papel pionero de las Universidades de la Iglesia, que (a contracorriente del recelo oficial) crean la Facultad de Sociología de Deusto (1963) y la de Ciencias Sociales León XIII (1964). La Facultad de Ciencias Políticas y Sociología de la Universidad Complutense no se crea hasta 1973, fruto de su escisión de la Facultad de Económicas, aunque con un número de alumnos bastante inferior (885 alumnos frente a los 19.811 estudiantes de económicas en aquel año).

La Iglesia católica también fue pionera en el impulso de investigaciones sociológicas. Principalmente, a través de fundaciones, como la Fundación FOESSA (Fomento de Estudios Sociales y Sociología Aplicada). De esta fundación partieron las ayudas a la investigación de los estudios sociológicos conocidos como *Informes FOESSA.* El primero de estos *Informes* se realiza en 1966, bajo la dirección de Amando de Miguel (al igual que el segundo, que data de 1970), que contó con la colaboración de un grupo de profesionales, que posteriormente impulsarían las tareas de investigación tanto en el sector público, como en el privado. En palabras de Lamo (1992: 127-128) estos *Informes* (refiriéndose a los dos primeros) representan "la prueba de madurez de la sociología española".

El paulatino afianzamiento de la sociología española encuentra también asideros en los primeros institutos de investigación de mercados y sociología aplicada, donde aparecen algunos nombres que, posteriormente, alcanzarán prestigio como figuras académicas. Son los casos de DATA y Amando de Miguel, o ECO y Jesús Ibáñez. Es el período, también, de aparición de centros privados como CEISA (Centro de Enseñanza e Investigación Sociológica), posteriormente *Escuela Crítica de Ciencias Sociales* (Ibáñez, 1992).

A partir de los años ochenta, la sociología española se adentra en una etapa de profesionalización y especialización, de reforzamiento institucional (en términos de creación de facultades o plazas universitarias), y de *investigación empírica.*

1.3. El contexto actual: diagnóstico y previsiones

En la actualidad, el avance de la sociología, tanto *empírica* como *teórica,* se caracteriza por la *síntesis* e *integración* de enfoques *macro* y *microsociológicos* en un mismo marco teórico y proceso de investigación.

A la pérdida de hegemonía del *paradigma funcionalista* (positivista, holista y nomológico), en la década de los sesenta, sucede, años más tarde, la crisis de otro monismo fundamental en la teoría social: el *marxismo.* A su descrédito como forma alternativa de ciencia social (por su jerarquización de niveles estructurales, su monocausalidad, su

incapacidad para comprender los aspectos culturales, emotivos y prerracionales del comportamiento social), se suma el derrumbe de los regímenes socialistas en el mundo, y el consiguiente desprestigio del *marxismo* como *paradigma de crítica política y social* (Pérez Agote, 1992; Giménez, 1992). La consecuencia de ambas crisis será el resurgimiento de las *tesis weberianas* en diferentes *enfoques subjetivistas* de la realidad social, a los que ya se ha hecho alusión (fenomenología, etnometodología, interaccionismo simbólico).

A partir de entonces, se suceden diferentes debates que enfrentan a las dos dimensiones básicas del análisis de la realidad social: la *estructural* y la *intersubjetiva*. Las viejas controversias entre *macro-microsociología, objetividad-subjetividad, explicación-comprensión, cuantitativo-cualitativo,* reaparecen, aunque ahora a la luz del nuevo enfoque analítico que supone la obra de Kuhn *La estructura de las revoluciones científicas* de 1962.

En los últimos años, se ha pasado de la controversia al acuerdo sobre la *pluralidad* de la realidad social. Su análisis debe abordarse desde diversos enfoques analíticos y metodológicos, si se pretende abarcar toda su complejidad.

La revolución de la *macro* y *microinformática,* con el consiguiente desarrollo y difusión de paquetes para el tratamiento de datos (ahora disponibles en grandes archivos nacionales e internacionales a un menor coste) ha propiciado la magnitud temporal y espacial, así como la sofisticación, del análisis de datos, tanto *cuantitativo* como *cualitativo*.

Las *líneas de investigación* más novedosas se centran, en la actualidad, en la aplicación de las *perspectivas estructurales* y de *redes* al comportamiento político. Asimismo, se consolidan los estudios de sociología del género, de la vejez, de la familia, del medio ambiente, y de las nuevas tecnologías (por citar algunas especialidades). Con estos estudios se pretende (como sucediera en los orígenes de la sociología) dar respuesta a los problemas generados por la transformación de la sociedad actual. Pero, con la gran diferencia de que ahora se dispone de más medios y mayor especialización en la obtención y análisis de la información (Rodríguez, 1991).

Lecturas complementarias

Caplow, T. (1977): *La investigación sociológica*. Barcelona, Laia.
García Ferrando, M. (1979): *Sobre el método*. Madrid, CIS.
Ibáñez, J. –comp.– (1992): *Las ciencias sociales en España. Historia inmediata, crítica y perspectivas*. Madrid, Universidad Complutense.
Kon, I. –comp.– (1989): *Historia de la sociología del s.XIX-XX*. Moscú, Progreso.
Rodríguez, J. A. (1991): "Nuevas tendencias en la investigación sociológica". *REIS*, 56, pp. 203-217.
Rodríguez Ibáñez, E. (1989): *La perspectiva sociológica. Historia, teoría y método*. Madrid, Taurus.

Ejercicios Propuestos

1. ¿Qué tres estudios destacaría por su contribución al desarrollo de la investigación social? Justifique la respuesta.
2. ¿Cuál es la principal aportación metodológica pionera de Weber al análisis de la realidad social? ¿En qué difiere de los planteamientos de su época?
3. Represente, de forma gráfica, la evolución de la metodología cuantitativa y la cualitativa, indicando las investigaciones más características de cada período en su desarrollo.
4. ¿Cuántos informes FOESSA se han efectuado hasta la fecha? Describa brevemente sus contenidos, destacando las divergencias y similitudes existentes entre ellos.

2

EL ANÁLISIS DE LA REALIDAD SOCIAL: APROXIMACIONES METODOLÓGICAS

La década de los años sesenta supuso –como ya se vio en el capítulo primero– el final de la hegemonía del *paradigma cuantitativo,* o *positivista,* y el renacer del *paradigma interpretativista* o *cualitativo.* A ello contribuyó la reaparición de las ideas fenomenológicas, junto con la nueva visión de la ciencia y del desarrollo del conocimiento científico propiciado por alumnos de Popper. Entre ellos se hallan Lakatos y Kuhn. Estos autores contribuyeron a la reafirmación del pluralismo teórico-metodológico en las ciencias sociales.

La fecha de 1967 fue clave en este resurgir (Tesch, 1992). En ese año se publicaron obras de gran repercusión en el desarrollo posterior de la sociología. Entre ellas destacan *Studies in ethnometodology,* de Garfinkel, o *The discovery of grounded theory,* de Glaser y Strauss. Un año antes (en 1966) había aparecido la obra de Bruyn, *Human perspective in sociology,* y la de Webb, Campbell, Schwartz y Sechrest, *Unobstrusive measures: nonreactive research.* Dos años después (en 1969), se edita el libro de Blumer, *Symbolic interactionism;* si bien, este último autor ya había publicado anteriormente (en 1967, también) un artículo sobre el mismo tema.

A partir de entonces, se multiplican las publicaciones sobre *metodología cualitativa.* Con ello se trata de llenar los huecos descubiertos en el *paradigma cuantitativo.* Este muestra ahora ser insuficiente para cubrir todas las demandas explicativas que exige el análisis de la realidad social.

2.1. El objeto y el método de la investigación social

En el desarrollo de las ciencias sociales han coexistido diferentes configuraciones, tanto sobre el *objeto* de estudio como sobre el *método* de análisis más adecuado. No son uno, sino varios los *objetos* que demarcan el área de la sociología. Berger y Luckman (1968) los resumen en tres:

a) La sociedad con su dinámica propia.
b) La sociedad como producto del hombre.
c) El hombre como producto de la sociedad.

Dependiendo de qué aspecto se anteponga, así será la explicación que resulte de la realidad social: positivismo, funcionalismo, teoría crítica, fenomenología, etnometodología, interaccionismo simbólico, teoría de sistemas.

Esta concepción plural del *objeto* conlleva diversas alternativas metodológicas, en la necesidad de adecuar el *método* al *objeto*. Como afirma Jiménez Blanco (1978: VII), "el método en una ciencia diría el mundo de objetos al que ésta se refiere", ya que ambos (el *objeto* y el *método)* son interdependientes.

2.1.1. Pluralidad metodológica

En la diversidad existente, se distingue una dicotomía metodológica que resulta "básica y radical en sociología" (Alvira, 1983: 54):

a) La "perspectiva *humanista/cualitativa",* que concibe le especificidad de las ciencias sociales. Rechaza el modelo metodológico de las ciencias naturales y aboga, en cambio, por el análisis de lo individual y concreto, por medio de la *comprensión* o interpretación de los significados intersubjetivos de la acción social (desde el punto de vista del actor social). El énfasis se pone en el *lenguaje* y en los aspectos *micro* de la vida social (situaciones cara a cara).
b) La "perspectiva *cientifista/cuantitativa".* Defiende la existencia de un único *método* (el de las ciencias naturales y exactas), general a todas las ciencias; al igual que el principio de *causalidad* y la formulación de *leyes generales* en el análisis de la realidad social. El énfasis se pone en la *"explicación,* en la contrastación empírica y en la medición objetiva de los fenómenos sociales".

Además del *cuantitativo* y *cualitativo,* Beltrán (1985) añade tres métodos más: el *histórico*, el *comparativo* y el *crítico-racional.* Además, subraya que la adecuación de cada uno depende de cuál sea el *objeto* de conocimiento en cada caso concreto.

El *método comparativo* y el *histórico* ya fueron destacados por el positivista Comte en 1839, en su *Curso de filosofía positiva,* como específico de la nueva ciencia denominada "Sociología", junto con la *observación* y el *experimento.* Asimismo, en 1895 Durkheim afirma, en *Las reglas del método sociológico*, que el método *comparativo* es "el único que conviene a la sociología (Durkheim, 1974: 133). No obstante, el *método histórico* adquirirá mayor relieve en la *dimensión intersubjetiva* del conocimiento sociológico.

Por su parte, el *método crítico-racional* fue resaltado por los autores englobados en la denominada *teoría crítica.* Estos autores defienden la consideración de los *valores*

y los *fines sociales* como *objeto* de estudio en las ciencias sociales. A ellos se suman la *comprensión* de los *significados* y la *interpretación* o *hermenéutica*.

García Ferrando (1979) realiza una tipología en función de qué criterio de clasificación se adopte. Si el criterio es la formación del conocimiento, diferencia el *método inductivo* del *deductivo*. En cambio, si la referencia constituye un área concreta de conocimiento, este autor diferencia el *método etnológico del psicoanalítico* y del *histórico*, entre otros.

Ibáñez (1985a), por el contrario, diferencia tres *perspectivas* en el análisis de la realidad social:

a) La *perspectiva distributiva*, cuyo ejemplo más general y concreto es la *encuesta social*.
b) La *perspectiva estructural*, cuyo ejemplo más general y concreto es el *grupo de discusión*.
c) La *perspectiva dialéctica*, que aplica también el *componente semiótico*, siendo el *socioanálisis* el ejemplo más general y concreto.

Años más tarde, Ortí (1989) modifica esta tipología metodológica, englobando la *perspectiva estructural* y la *dialéctica* en una misma perspectiva, y separada de la *distributiva*. De esta forma, la clasificación de Ortí se asemeja a la anteriormente señalada como básica: la *perspectiva cuantitativa* y la *cualitativa*.

a) El *enfoque distributivo* se corresponde con el método de las ciencias naturales, con la *explicación* y con los *métodos cuantitativos*.
b) El *enfoque estructural-dialéctico* asume la especificidad de las ciencias sociales, así como la *comprensión* y los *métodos cualitativos*.

En conclusión, se distinguen dos *perspectivas metodológicas fundamentales:* la *cuantitativa* y la *cualitativa*. Ambas difieren tanto en la estrategia seguida en la recogida de información como en su análisis, debido a su vinculación a distintas perspectivas paradigmáticas. Pero, dado que ambas proporcionan un marco filosófico y metodológico concreto para el estudio de la realidad social, aquí se ha optado –siguiendo a Filstead (1986)– por conceptualizarlas como *paradigmas*.

Un *paradigma* representa un modelo fundamental, "una *imagen* básica del objeto de una ciencia. Sirve para definir lo que debe estudiarse, las preguntas que es necesario responder, cómo deben preguntarse y qué reglas es preciso seguir para interpretar las respuestas obtenidas" (Ritzer, 1993: 598). Kuhn (1978) utilizó precisamente el término *paradigma* para referirse a este tipo de acuerdo entre *métodos* y *procedimientos*.

En el Cuadro 2.1 se resumen algunas de las características distintivas que tradicionalmente se han atribuido a los dos *paradigmas* en cuestión.

CUADRO 2.1. Características diferenciadoras de los paradigmas cuantitativo y cualitativo.

	Paradigma cuantitativo	*Paradigma cualitativo*
Base epistemológica	Positivismo, funcionalismo.	Historicismo, fenomenología, interaccionismo simbólico.
Énfasis	*Medición* objetiva (de los hechos sociales, opiniones o actitudes individuales); demostración de la *causalidad* y la gene- *ralización* de los resultados de la investigación.	El actor individual: *descripción* y *comprensión interpretativa* de la conducta humana, en el propio marco de referencia del individuo o grupo social que actúa.
Recogida de información	*Estructurada* y sistemática.	*Flexible:* un proceso interactivo continuo, marcado por el desarrollo de la investigación.
Análisis	*Estadístico,* para cuantificar la realidad social, las relaciones causales y su intensidad.	*Interpretacional,* socio-lingüístico y semiológico de los discursos, acciones y estructuras latentes.
Alcance de resultados	*Nomotética:* búsqueda cuantitativa de leyes generales de la conducta.	*Ideográfica:* búsqueda cualitativa de significados de la acción humana.

El *paradigma cuantitativo* corresponde a la *epistemología positivista,* a la aquí denominada *dimensión estructural* del análisis de la realidad social. Por el contrario, el *paradigma cualitativo* se asocia a la *epistemología interpretativa (dimensión intersubjetiva),* centrada en el *sujeto* individual y en el descubrimiento del *significado,* los *motivos* y las *intenciones* de su *acción.*

El concepto de *"verstehen"* de Weber se ajusta a esta última dimensión: el investigador debe procurar analizar el mundo social a través del sujeto que actúa, con objeto de entender los *significados,* los *motivos* e *intenciones* de su *acción.* Pero, a veces, estos planteamientos se tergiversan. Como apunta Brannen (1994: 6), "mucha de la investigación cualitativa es simplemente descriptiva". Igualmente, la *investigación cuantitativa* no siempre se enfoca a la *comprobación de hipótesis* (deducidas de *teorías* existentes); sino que, con frecuencia, se dirige sólo a la *descripción* de hechos sociales, opiniones o actitudes individuales, sin ninguna intención explicativa.

Asimismo, aunque sea cierto que la *investigación cualitativa* busque entender el punto de vista de la persona analizada, es –en palabras de Hammersley (1994: 45)– "muy raro que la investigación cualitativa se restrinja a documentar el innato punto de vista".

Por estas y otras razones, hay que insistir en el carácter genérico de los atributos que se hallan resumidos en el Cuadro 2.1 (no exentos de controversia). Además, ha de subrayarse –como hacen O'Connell y Layder (1994: 32)– que "aunque existen metodólogos que se adhieren a los extremos de o la filosofía positiva o a la interpretativa, existen también muchos que caen en el medio de este *continuum* metodológico y filosófico".

En enfrentamiento entre ambos *paradigmas* marcó el desarrollo de la sociología en los años setenta. La inexistencia de un *paradigma* dominante, con el consiguiente enfrentamiento de los enfoques teórico-metodológicos existentes (en busca de hegemonía o supremacía sobre los demás), provoca la *crisis* de la sociología y la consiguiente redefinición de su objeto de conocimiento.

A partir de entonces, los esfuerzos de teóricos y metodólogos se dirigen a buscar una solución a la situación de *crisis,* de pérdida de rumbo, de la sociología.

Se reconoce la pluralidad de vías para acceder a la realidad social, pero no se trata de afirmar un *paradigma* sobre otro, sino de buscar compatibilidades entre ellos. La frase de Reichardt y Cook (1979: 27): "Es tiempo de dejar de construir muros entre métodos y empezar a construir puentes", ejemplifica esta predisposición a la *complementariedad* entre *paradigmas,* y no a su oposición.

Como consecuencia, un nuevo concepto se pone de moda, sobre todo en los años ochenta: la *triangulación* o utilización de distintos métodos en una misma investigación.

2.1.2. Articulación y convergencia interparadigmática: la triangulación

El concepto de *triangulación* tiene sus raíces en otros ámbitos de conocimiento diferentes al de la investigación social. Concretamente, en la topografía, en la navegación e, incluso, en la logística castrense. En todos estos ámbitos, el término de *triangulación* adquiere un mismo significado: la utilización de múltiples puntos de referencia para localizar la posición exacta de un objeto en el espacio. De esta forma se logra una mayor precisión que la alcanzada mediante la aplicación de un único punto de referencia.

En las ciencias sociales, el término *triangulación* adquiere un significado similar. Por él se entiende la aplicación de distintas metodologías en el análisis de una misma realidad social.

Esta *estrategia de investigación* no es novedosa. Se remonta a los mismos orígenes de la investigación social. A autores como Booth, quien, por ejemplo, en su investigación *Vida y trabajo de los habitantes de Londres (1889-1891)*, compagina la utilización de *datos disponibles* (censos de población, informes de inspectores sanitarios y de escuelas, datos policiales, de organizaciones de beneficencia) con *observación directa* en los propios barrios, y *entrevistas* a hombres de negocios.

Pero la necesidad de articular distintos métodos en una misma investigación no se plantea a fondo hasta los años cincuenta, cuando se debate la conveniencia de combinar la *encuesta* con el *trabajo de campo* (Vidich y Shapiro, 1955; Zelditch, 1962; Mc-

Call y Simmons, 1969; Sieber, 1973); o técnicas concretas de recogida de información como la *entrevista* y la *observación participante* (Becker y Geer, 1957; Trow, 1957).

A finales de los cincuenta, en 1959, Campbell y Fiske escriben un artículo, que iba a tener una gran repercusión, en la revista *Psychological Bulletin* ("Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix"). En este artículo se desarrolla la idea de la *operacionalización múltiple:* la necesidad de utilizar más de un método en el proceso de medición para, de esta forma, aumentar la *validez* de los hallazgos y el grado de confianza en los mismos.

Por *validez* se entiende el grado de adecuación de una medida particular respecto a la realidad a la que apunta. En opinión de Campbell y Fiske (1959), la adecuación de cualquier *constructo hipotético* debe establecerse por su:

a) *Validez convergente:* el acuerdo entre los distintos métodos de medir el mismo constructo.
b) *Validez discriminante:* la diferenciación entre los resultados cuando se emplea el mismo método para asegurar cualidades hipotéticamente diferentes de los mismos objetos.

En 1966, Webb y otros autores (entre los cuales figura Campbell) publican su obra *Unobstrusive measures: nonreactive research in the social sciences,* donde por vez primera aparece el término *triangulación.* Basándose en las ideas de Campbell y Fiske (1959), estos autores (Webb *et al.,* 1966: 3) defienden la necesidad de aplicar una "*triangulación* de los procesos de medición". Con ello se persigue el logro de un mayor grado de *validez* de los hallazgos de una investigación.

Unos años más tarde, en 1970, Denzin dedica un apartado de su obra *The Research Act* (la Parte IV concretamente) a la definición de la *triangulación* en la investigación social. A partir de entonces se suceden las publicaciones que tratan la *triangulación.* Entre ellas se halla el artículo de Jick, "Mixing Qualitative and Quantitative Methods: Triangulation in Action", de 1979; o, el magnífico libro de Brewer y Hunter, *Multimethods research: a synthesis of styles,* de 1989, entre otras publicaciones.

Los argumentos de Denzin a favor de la *triangulación* alcanzan una gran repercusión en diferentes campos de conocimiento, como en la *evaluación de programas* (Reichardt y Cook, 1979; Schwartz y Jacobs, 1979; Patton, 1980; Madey, 1982) o en la *etnografía* (con las aportaciones de Hammersley y Atkinson, 1983; Burgess, 1984, y Hammersley, 1992, entre otros). Si bien, aparecen también detractores (como Silverman, 1985; Fielding y Fielding, 1986; o Blaikie, 1991). Estos últimos autores reconocen que la *triangulación* favorece la adquisición de un conocimiento más amplio y profundo de la realidad social, aunque niegan que este conocimiento sea necesariamente más objetivo y fiable que el obtenido mediante una única *metodología.*

Pero, ¿en qué consiste la *triangulación?* Siguiendo a Denzin (1975), existen cuatro tipos básicos de *triangulación* posibles:

a) Triangulación de datos.
b) Triangulación de investigadores.
c) Triangulación teórica.
d) Triangulación metodológica (intramétodo y entre métodos).

• *Triangulación de datos*

Este tipo de *triangulación* es el más habitual en la práctica de la investigación social. Consiste en la utilización de varias y variadas fuentes de información sobre un mismo objeto de conocimiento, con el propósito de contrastar la información recabada.

EJEMPLO DE TRIANGULACIÓN DE DATOS

Para ilustrar la *triangulación de datos*, a continuación se expone su materialización en una investigación concreta: una *investigación evaluativa* que, hace unos años, efectué sobre la actuación de los Tribunales Tutelares de Menores (Cea, 1992). En este estudio se acudió a tres fuentes de información diferentes para la medición de un mismo *concepto teórico:* el de *reincidencia* en conducta delictiva, una vez concluida la intervención del Tribunal.

a) La información contenida en el expediente del menor. Concretamente, las *detenciones* policiales o de la Guardia Civil registradas, y las *denuncias* efectuadas por familiares u otras personas.
b) La proporcionada por la Dirección General de la Policía, concerniente a *detenciones* policiales y de la Guardia Civil, posteriores a la intervención del Tribunal Tutelar.
c) La facilitada por la Dirección General de Instituciones Penitenciarias, relativa a si el menor había estado, o se hallaba en el momento de la investigación, en prisión y cuál era su situación procesal.

Con la contrastación de estas tres fuentes de información se perseguía analizar la coincidencia o divergencia en los datos obtenidos; y, en caso de divergencia, a qué se debía ésta. Lógicamente, cuando la información coincide, la evidencia adquiere una fuerza mayor.

• *Triangulación de investigadores*

Este tipo de *triangulación* sería el equivalente a lo que se conoce, en la actualidad, por *equipos interdisciplinarios:* la realización de una misma investigación por un

equipo de investigadores (procedentes de distintas áreas de conocimiento o especialistas en diferentes *metodologías)*, que observan un mismo objeto de estudio desde diferentes puntos de vista, en función de la disciplina científica a la que pertenezcan.

• *Triangulación teórica*

Este es el tipo básico de *triangulación* aunque, en opinión de Denzin (1975), es el menos alcanzable en la práctica de la investigación social.

Implica considerar todas las hipótesis que puedan extraerse de un mismo problema de investigación. Ello implica abarcar el mayor número de *perspectivas de análisis* posible. La actuación en una *triangulación teórica* se resume en la Figura 2.1.

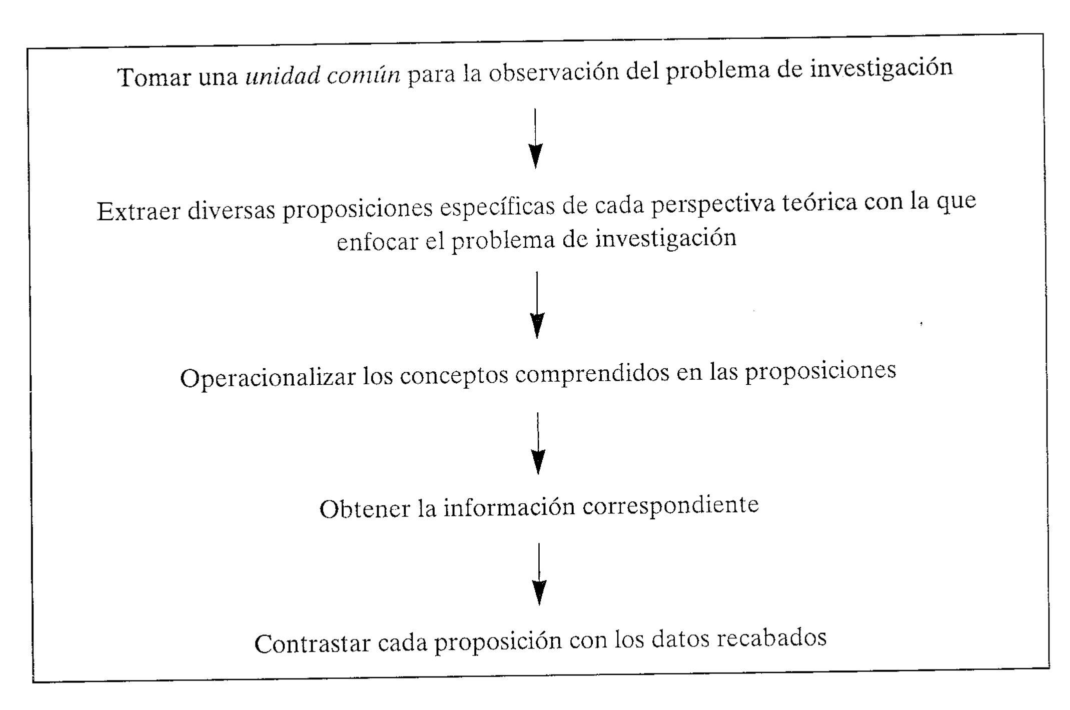


Figura 2.1. El procedimiento para una *triangulación teórica.*

Como es probable que cada perspectiva demande un tipo diferente de información, se aplicaría, a su vez, la *triangulación de datos*.

La *triangulación teórica* también puede implicar el empleo de diversos investigadores para evitar la posibilidad de sesgos teóricos explícitos que, de otra forma, podrían introducirse.

EJEMPLO DE TRIANGULACIÓN TEÓRICA

En una investigación que tuviese como objetivo principal el estudio de los factores que inciden en la comisión de conducta delictiva, el investigador tendría que partir –para el cumplimiento de la *triangulación teórica*– de las diversas perspectivas teóricas existentes al respecto: tanto aquellas que enfatizan el nivel *macroestructural*, como las que indagan en los procesos *microinterpersonales;* las teorías que explican la causalidad de la *primera transgresión*, en conjunción con las que analizan la influencia de la aplicación de la norma legal en el desarrollo de la *carrera delictiva*. Mediante esta articulación teórica podría alcanzarse una visión completa de la causalidad de la delincuencia.

- *Triangulación metodológica*

Representa la última forma genérica de *triangulación*. Comprende las dos modalidades siguientes:

1) La triangulación intramétodo.
2) La triangulación entre métodos.

1) *Triangulación metodológica intramétodo*

Este tipo de *triangulación metodológica* se aplica cuando el investigador escoge un único método o estrategia de investigación, pero aplica distintas técnicas de recogida y de análisis de datos; o cuando repite el mismo método en situaciones y momentos diferentes. El objetivo que se persigue es, igualmente, comprobar la validez y la fiabilidad de la información que primeramente se ha obtenido.

EJEMPLO DE TRIANGULACIÓN METODOLÓGICA INTRAMÉTODO

En la investigación evaluativa anteriormente referida (Cea, 1992) también se puso en práctica la *triangulación metodológica intramétodo*. Consistió en la aplicación de tres técnicas de análisis multivariable para el cumplimiento de un mismo objetivo: la descripción del menor de reforma. Los datos se obtuvieron de una única fuente de información: el expediente del menor incoado por el Tribunal Tutelar de Menores de Madrid. Si bien, para su análisis, se acudió a las técnicas multivariables siguientes:

1) *Análisis de conglomerados*, para comprobar si existen o no diferentes tipos de menores de reforma, su peso e importancia en el conjunto de la población de reforma.

2) *Análisis discriminante,* con objeto de comprobar si los tres grupos de menores obtenidos (mediante la técnica analítica anterior) en realidad quedaban bien definidos por las variables apuntadas.
3) *Análisis factorial confirmatorio,* para corroborar los resultados alcanzados con las técnicas multivariables anteriores. Con su aplicación se obtienen las variables que más significativamente identifican a cada grupo de menores.

El inconveniente que Denzin (1975) encuentra en este tipo de *triangulación* es que únicamente se emplea un solo *método*. Por lo que, no se logra superar las limitaciones propias del *método* de investigación escogido.

Así, por *ejemplo,* la investigación antes comentada habría logrado una más amplia y profunda descripción del menor de reforma, si la información analizada no se hubiese limitado a la registrada en el expediente del menor. La información extraída del expediente (relativa a las características sociodemográficas y legales del menor) podría haberse contrastado, o complementado, con la facilitada por otras fuentes: datos procedentes de otros estudios similares –siempre y cuando los datos fueran comparables; o información obtenida mediante *estudio de casos,* con *entrevistas en profundidad* a menores equiparables, a personas de su entorno sociofamiliar u otras, por citar otras fuentes de información.

2) *Triangulación metodológica entre métodos*

A diferencia de la modalidad anterior, la *triangulación entre métodos* es "la más satisfactoria" (Denzin, 1975) y "la más popular" (Jick, 1979).

Consiste en la combinación de *métodos* de investigación (no similares) en la medición de una misma unidad de análisis. Con ello se pretenden paliar las limitaciones de cada *método*, contrarrestándolas con las potencialidades de los otros *métodos*. De esta forma, la *triangulación entre métodos* se convierte en vehículo para la *validación cruzada* (Jick, 1979), cuando se alcanzan los mismos resultados con métodos distintos, pero que apuntan a la misma dimensión del problema de investigación.

EJEMPLO DE TRIANGULACIÓN METODOLÓGICA ENTRE MÉTODOS

Imaginemos que se desea estudiar la violencia en las calles. El equipo investigador, sin duda, lograría una mejor descripción de su objeto de estudio, si se decidiese por la articulación de varias *estrategias de investigación,* haciendo uso de la *triangulación metodológica entre métodos.* Ello le llevaría a la conjugación de datos procedentes de fuentes diversas.

Una de las posibles combinaciones podría ser la siguiente: adicionar *datos estadísticos* con *datos etnográficos.*

- Los *datos estadísticos* procederían, asimismo, de distintas fuentes: las *Estadísticas Judiciales* del INE; las estadísticas proporcionadas por la Dirección General de la Policía; y las propias de la Dirección General de Instituciones Penitenciarias.
- Los *datos etnográficos* se obtendrían mediante las técnicas de la *observación* (en lugares habituales donde acontecen actos de violencia) y de la *entrevista*. Las *entrevistas abiertas* se harían tanto a los hacedores de actos de violencia como a las víctimas.

Con estas cuatro modalidades genéricas de *triangulación* pueden efectuarse una variedad de combinaciones. Denzin (1975) defiende la *triangulación multimétodo* como ideal. Esta consistiría en la articulación, en un mismo estudio, de los cuatro tipos de *triangulación: de datos, investigadores, teórica y metodológica. De esta forma se alcanzarían* unos resultados más satisfactorios que si, únicamente, se hubiera triangulado por fuentes de *datos, investigadores, teorías* o *métodos*.

Al discurso de Denzin (1975), hay que añadir las precisiones enunciadas por Brewer y Hunter (1989) para el ejercicio, con éxito, de la *triangulación:*

a) Una *triangulación* exitosa requiere un análisis cuidadoso de cada *método,* en relación a los otros *métodos,* y también respecto a las demandas del proceso de investigación.

b) Qué *métodos* y cuántos *métodos* han de emplearse depende de la información que se precise para esclarecer el problema que se investigue.

c) Muchos estudios *multimétodo* tienen objetivos mezclados. Se puede generar una teoría con un *método* y comprobarla con otro.

d) Tanto al crear como al comprobar una *teoría,* ciertas variables pueden ser accesibles a un *método,* pero no a otros. Por lo que se precisa una serie de *métodos* de recogida de datos, bastante vinculados, para considerar la significatividad de diferentes tipos de variables.

e) Hay que procurar que los distintos métodos que se apliquen para obtener información midan realmente los mismos *conceptos teóricos*. Ello es necesario, si se pretende que los resultados obtenidos (mediante diferentes procedimientos) sean comparables.

f) Para que la aplicación de un *método* no influya en los resultados alcanzados por otro *método,* sería conveniente seguir alguna de las opciones siguientes:

1) Realizar la *recogida de datos escalonadamente,* de manera que la información que sea más vulnerable a la influencia del investigador o de los sujetos analizados (generalmente aquélla obtenida mediante *métodos reactivos,* como la *encuesta)* sea recabada primero.

En general, cuando se conjugan *estrategias de investigación* que provocan un grado diferente de reactividad en los sujetos que se analizan, se obtienen mejores resultados, si las *estrategias* que generan mayor *reactividad* (caso de la *encuesta)* se aplican con anterioridad a las *"no reactivas"* (como el *uso de fuentes documentales y estadísticas).*

EJEMPLO DE RECOGIDA DE INFORMACIÓN DE FORMA ESCALONADA

Se quiere analizar el funcionamiento de una institución concreta, mediante la articulación de dos *estrategias metodológicas*: el *uso de fuentes documentales y estadísticas* (los informes de la institución) y una *encuesta*. Para evitar la influencia que la práctica de una de las *estrategias* pudiera tener en los resultados de la otra, la recogida de información podría efectuarse de manera escalonada.

Primero, se aplicaría la *encuesta* y, después, se accedería a los *informes* del centro. Así se evitaría el efecto que el conocimiento del acceso (por parte del equipo investigador) a los *informes* del centro pudiera tener en las respuestas de los sujetos a la *encuesta*.

2) *Dividir la muestra en submuestras,* o *extraer varias muestras,* y en cada una de ellas, aplicar una técnica distinta de recogida de información. Ello evitaría, igualmente, el efecto directo de un *método* sobre otro, al no exponerse a los mismos sujetos a más de un *método*.

   Asimismo, podrían comprobarse los efectos concretos de un *método* en otro. Para esto último, se compararían los resultados obtenidos mediante este procedimiento con aquellos alcanzados después de exponer a algunos sujetos (extraídos de las *muestras* de forma *aleatoria)* a ambos tratamientos.

   Pero, téngase presente que este procedimiento de extracción de varias *muestras* no garantiza la inmunidad frente a los *sesgos* derivados del conocimiento, por parte de los integrantes de una de las muestras, de las experiencias de investigación de las otras *muestras*. De ahí que, si se desea evitar este tipo de *sesgo,* sea necesario, o bien aplicar cada una de las técnicas de recogida de datos escalonadamente, o bien proceder de forma rápida y simultánea en las distintas *muestras*.

En conclusión, si la finalidad que se persigue mediante la *triangulación* es comprobar la *validez* de los resultados de la investigación, la clave se halla en garantizar la independencia en todo el proceso. Quiere esto decir, que las distintas *estrategias de investigación* han de practicarse independientemente unas de otras; pero dirigidas a la medición de los mismos *conceptos teóricos* por vías diferentes.

Por el contrario, el requisito de la independencia no es imprescindible, si el objetivo del estudio es la exploración: una primera aproximación a fenómenos poco co-

nocidos, con la finalidad de identificar variables relevantes y, a partir de ellas, extraer *hipótesis* para comprobar en estudios posteriores.

• Respecto a la *aplicación de la triangulación,* quizás los intentos más destacados se han concentrado en la integración del *trabajo de campo* y la *encuesta* (Jick, 1979). Desde distintos ámbitos se ha defendido la necesidad, y viabilidad, de dicha articulación.

Vidich y Shapiro fueron de los primeros en exponer la necesidad de integrar el *trabajo de campo* con la *encuesta,* aunque conceden un mayor protagonismo a la *encuesta.* La *encuesta* contribuye a la generalización de los hallazgos del *trabajo de campo* cualitativo. "Sin los datos de *encuesta*, el *observador* sólo puede realizar conjeturas sobre su área de ignorancia" (Vidich y Shapiro, 1955: 31).

Años más tarde, en 1973, Sieber demuestra en un magnífico artículo –publicado en la revista *American Journal of Sociology* ("The integration of fieldwork and survey methods")– los beneficios mutuos de integrar el *trabajo de campo* y la *encuesta,* tanto en las *fases exploratorias* de la investigación como en las de *verificación.*

Entre las contribuciones del *trabajo de campo* (mediante observación participante, entrevistas informales, y el uso de datos disponibles) al análisis e interpretación de los datos de *encuesta,* este autor destaca las siguientes:

a) La *estructura teórica* que guía el análisis de *encuesta* puede derivarse, total o plenamente, del *trabajo de campo* cualitativo. Ello exige que el *trabajo de campo* cualitativo preceda (en el diseño de la investigación) a la *encuesta.*
b) Los resultados de la *encuesta* pueden ser validados mediante el recurso a *observaciones* y *entrevistas informales.* De manera especial, cuando los resultados que se obtengan de la *encuesta* sean inesperados o sorprendentes; o cuando la *encuesta* obtenga un bajo porcentaje de respuestas.

   Para que el *trabajo de campo* cumpla la función de aclaración de los resultados de la *encuesta,* la *encuesta* deberá precederle en el tiempo.
c) Las relaciones estadísticas, obtenidas con la *encuesta,* pueden interpretarse mediante referencias a *observaciones* y *entrevistas informales.*
d) El *trabajo de campo* también ayuda a la identificación de los sujetos más adecuados para ser encuestados.

   A veces existen sujetos especiales, cuya existencia se descubre gracias a que previamente se ha realizado un *trabajo de campo exploratorio.* Ello posibilita su inclusión en el *diseño de la muestra* a encuestar.
e) El *trabajo de campo* contribuye, igualmente, a la elaboración del *cuestionario:*

   1) En la formulación de los *ítems* que lo conforman.
   2) Clarificando las preguntas confusas de un *cuestionario.*

Ibáñez (1989) aconseja efectuar una *investigación estructural* (generalmente algún *grupo de discusión)* como fase previa a la realización de una *encuesta.* La finalidad es evitar que la *encuesta* se convierta en la opinión de quien redacta el *cuestionario.*

Fuller *et al.* (1993) añaden otras contribuciones del *grupo de discusión* al *diseño del cuestionario:*

1) Ayuda a la identificación de los *conceptos teóricos* relevantes (y elimina los conceptos irrelevantes).
2) Colabora en la formulación de *hipótesis*.
3) Sugiere qué *lenguaje* utilizar para una mejor comunicación con la población del estudio.

EJEMPLO DE CONJUNCIÓN DE ENTREVISTAS ABIERTAS, GRUPOS DE DISCURSIÓN, Y ENCUESTA EN UNA MISMA INVESTIGACIÓN

En una investigación evaluativa sobre los Centros Municipales de la Tercera Edad (Cea y Valles, 1993), se hizo uso de la articulación susodicha. Primero, se acudió a *entrevistas abiertas* (o en *profundidad*) a profesionales responsables de los Centros Municipales de la Tercera Edad. La finalidad era obtener una visión conjunta del funcionamiento de los centros municipales que facilitase el diseño posterior de la *encuesta*: tipo de preguntas y forma de realizarlas.

Después de la *encuesta*, se efectuaron varios *grupos de discusión* (tanto con socios de los Centros Municipales de la Tercera Edad, como con no socios de la misma edad, por separado). El propósito ahora no era exploratorio ni descriptivo, sino explicativo: aclarar, y profundizar, en las conclusiones a las que se llegó tras el análisis estadístico de los datos de *encuesta*.

A la inversa, la *encuesta* también favorece –siguiendo a Sieber (1973)– al *trabajo de campo:*

a) *Identificando* tanto los casos relevantes, como los no relevantes (o no representativos), para ser después analizados mediante el *trabajo de campo*.
b) *Demostrando* la generalidad de las *observaciones de campo*.
c) *Verificando* las interpretaciones de campo.
d) *Aclarando* las observaciones "mal interpretadas" o "inexplicables". Por ejemplo, proporcionando información de informantes o sujetos que fueron previamente pasados por alto.

Las otras *estrategias de investigación (experimentación, uso de datos disponibles)* también se benefician de su aplicación complementaria con otras *estrategias*.

De la *experimentación* se afirma que resulta más eficaz si va precedida de un *estudio de casos,* que proporcione un conocimiento más amplio y preciso de la situación (French, 1972).

Por su parte, el empleo de *datos disponibles* puede servir de base (generación de *hipótesis),* o de fundamento, de los hallazgos obtenidos mediante otras *estrategias de investigación.*

En suma, todas las *estrategias de investigación* poseen una ventajas y unos inconvenientes. Mediante su articulación se intentan ajustar las potencialidades de cada una de ellas con cada uno de los objetivos marcados en la investigación. De esta forma, las limitaciones de cada *estrategia* se solventan con las potencialidades de las otras. Así, por ejemplo:

a) La *encuesta* aparece como la *estrategia de investigación* más adecuada, cuando se precisa información (hechos, opiniones, actitudes) de un gran número de personas, localizadas en una área geográfica amplia, en un breve período de tiempo, y a un precio relativamente bajo.
b) El *estudio de casos,* cuando se desea una información más profunda de la realidad social. Especialmente, de las relaciones sociales.
c) La *experimentación* se centra más en la comprobación de *hipótesis causales.*
d) El *uso de datos disponibles* se muestra más adecuado para indagar en hechos del pasado y procesos de cambio social. Si bien, en ambos casos se aconseja combinar la información proporcionada por esta *estrategia* con otras *estrategias de investigación.*

En conclusión, ambas *metodologías (cuantitativa* y *cualitativa)* se adecuan a cualquier fase del proceso de investigación social.

Tradicionalmente, el *paradigma cuantitativo* ha relegado al *cualitativo* al nivel que Reichenbach (1973) denomina "contexto de descubrimiento": se caracteriza por la realización de actividades *exploratorias,* que contribuyen a la generación de *hipótesis.* Por el contrario, los *métodos cuantitativos* se reservan el "contexto de la justificación" científica de los hallazgos, su significatividad. No obstante, la práctica de la investigación social demuestra que lo inverso también es cierto.

Desde su experiencia en la investigación social comparativa, Allardt (1990) argumenta que por medio de datos *cuantitativos* el investigador es capaz de explorar los fenómenos estudiados. Si bien, para que los resultados obtenidos sean más interpretables, es preciso que se efectúen estudios *cualitativos,* mediante los cuales pueda profundizarse en el tema de estudio.

De igual manera se pronuncian otros autores, como Walker. Para este último autor, "la investigación cualitativa, más generalmente, puede ayudar a interpretar, iluminar, ilustrar y cualificar las relaciones determinadas estadísticamente" (Walker, 1988: 22).

En consecuencia, ambas *metodologías* son intercambiables. De lo que se trata no es de subordinar unas *estrategias de investigación* a otras, sino de *articularlas* o *combinarlas* en un mismo proyecto de investigación, en función de cuál sea el objetivo concreto a cubrir. Como subraya Beltrán (1989):

CUADRO 2.2. Ventajas e inconvenientes de la *triangulación*.

Ventajas	*Inconvenientes*
Al cubrir más dimensiones de la realidad social, se alcanza una información más profunda y diversificada.	Puede suponer un mayor *coste* económico y temporal que aplicar un único método de investigación (1).
Los resultados del estudio adquieren un grado mayor de *validez* (si se llegan a los mismos hallazgos con métodos diferentes).	Dificultad de encontrar una *unidad de observación* común a la que aplicar distintos métodos
La *teoría* queda más reforzada, en caso de ser confirmada. Si no lo es, proporciona un fundamento mayor para su modificación.	Dificultad de hallar un equipo de investigadores bien formados en ambas *metodologías* (cuantitativa y cualitativa).
Facilita la *comparabilidad* de los datos por métodos distintos.	Mayor dificultad para repetir la investigación.
Demuestra que ambas *metodologías* (cuantitativa y cualitativa) pueden y deben utilizarse de forma complementaria.	El uso de una técnica de obtención de información puede afectar a observaciones posteriores que se efectúen mediante otras técnicas (2).
	Problemas de *comparabilidad,* debido a la obtención de diferentes porcentajes de respuesta (en cada una de las técnicas de obtención de datos utilizada); o por haber distintas fuentes de error en cada método, y no haberse detectado.
	Posibilidad de primar, sin justificar, un método sobre otro.
	Existencia de incompatibilidad epistemológica entre algunos métodos (3).

(1) Para solventar este inconveniente, Denzin (1975) aconseja utilizar más adecuadamente los medios disponibles. Por ejemplo, puede que no resulte posible acudir a varios investigadores, aunque sí emplear diversas fuentes de datos.

(2) Bien debido a que la experiencia alcanzada con el primer método lleve al investigador a utilizar el segundo método de forma diferente e, incluso, mejor que el primero; o bien debido a la influencia en los sujetos (a lo que ya se ha hecho referencia).

(3) La principal objeción a la combinación metodológica se expresa, con frecuencia, en términos de "pertinencia epistemológica". Autores como Blaikie (1991), consideran a ésta como la barrera fundamental para integrar formas incompatibles de conocimiento.

"Basta con afirmar el método cualitativo junto al cuantitativo, dejando que sea el propio objeto de conocimiento el que lo justifique y reclame en función de sus propias necesidades, perfectamente diferenciadas" (Beltrán, 1989: 40).

La *complementariedad paradigmática* se impone como "una necesidad práctica impuesta por la investigación" (Pérez Agote, 1989: 118).

La *triangulación de estrategias de investigación* presenta unas ventajas, pero también, unos inconvenientes. Ambos se resumen en el Cuadro 2.2.

2.1.3. Ejemplos de la práctica de la triangulación en investigaciones reales

Como ejemplos ilustrativos, a continuación se hace una breve referencia a cuatro investigaciones reales en las que se efectúa alguna, o varias, de las modalidades de *triangulación* comentadas. Las dos primeras investigaciones se desarrollan en Gran Bretaña; las dos últimas, en España.

La inclusión de dos investigaciones no realizadas en España –a diferencia de otros capítulos del presente manual– responde a su interés por constituir *investigaciones longitudinales* en las que la *triangulación* se cumplimenta en orden inverso.

- *Laurie, H. y O'Sullivan (1991)* Combining qualitative and quantitative data in the longitudinal study of household allocations. *Sociological Review, vol. 39, pp. 113-130.*

Esta primera investigación aún se halla en fase de ejecución, en el marco de la Universidad de Essex (ESRC). Se trata de un *estudio longitudinal,* cuyo objetivo principal es la medición del cambio, o la continuidad, en la distribución de recursos en hogares de diferente ubicación.

Consta de una *encuesta* (sobre una variedad de temas socioeconómicos), que se aplica a una muestra compuesta por 5.000 hogares. En ellos se entrevistará anualmente, desde 1991 hasta 1999, a todas las personas adultas.

En la fase preliminar de la investigación, se lleva a cabo un *estudio cualitativo,* cubriéndose 77 hogares. Dicho estudio consta de:

a) 19 *entrevistas en profundidad* a cada uno de los integrantes de la pareja (a los que se entrevista de forma separada).
b) 7 *grupos de discusión*, de 8 ó 10 personas cada uno.

El componente *cualitativo* queda, en consecuencia, limitado a su uso "convencional", como precursor de la *estrategia de investigación* fundamental: la *encuesta.* De la *encuesta,* por el contrario, se espera que proporcione una medición, de forma sistemática, de los *indicadores* que los *análisis cualitativos* destaquen como relevantes.

No obstante, los autores del estudio (Laurie y Sullivan) defienden la necesidad de recoger información *cualitativa* (de forma continua) como un componente fundamental de las *encuestas longitudinales* tradicionales.

Las *encuestas panel* permitirán la identificación de los acontecimientos que hayan acaecido durante el año proporcionando, de esta forma, datos valiosos sobre la secuencia temporal de los acontecimientos. En cambio, los datos *cualitativos* permitirán situar estos acontecimientos en el contexto y explorar su significado para los actores implicados.

- *Ritchie, J. (1990):* Thirty families. Their living standards in unemployment. *London, Department of Social Security Research Report, nº 1, H.M.S.O.*

Esta investigación también consiste en un *estudio longitudinal*, en el que se articula la *metodología cuantitativa* con la *cualitativa*, pero en orden inverso a la investigación anterior. Se trata de una investigación llevada a cabo por el Departamento de Seguridad Social británico, en la que se analiza la situación de desempleo y su repercusión en las formas de vida. Consta de:

a) Una *encuesta* a 3.000 personas entre 20 y 58 años de edad (cabezas de familia, de ambos sexos, que han estado, ininterrumpidamente, en situación de desempleo, durante un mínimo de tres meses, al principio del estudio). El trabajo de campo se desarrolla en los otoños de 1983 y 1984.

El objetivo fundamental del estudio es indagar el modo en que se modifican las formas de vida, tanto al principio como durante la situación de desempleo. Para ello se efectúa este *estudio de cohorte,* con entrevistas realizadas 3 y 15 meses después de que la persona desempleada empezase a registrarse como "parada".

b) Un pequeño número de *estudios de casos* entre personas en paro con familia. Con ello se pretende obtener una información diferente, pero complementaria a la *encuesta* (para esclarecer los resultados estadísticos).

El *estudio de casos* se lleva a cabo en dos fases. Una primera, en paralelo a la *encuesta;* y, la segunda fase, cuatro años después. Su interés principal es investigar el proceso que lleva a cualquier cambio en las formas de vida durante el desempleo, y el impacto que estos cambios provocan en las familias. La información se recoge mediante *entrevistas en profundidad.*

- *Arribas Macho, J. y González, J. (1987):* La Juventud de los ochenta. Estudio sociológico de la juventud de Castilla y León. *Junta de Castilla y León. Consejería de Educación y Ciencia.*

Una investigación sociológica en la que se analiza la situación, en los años ochenta, de los jóvenes de edades comprendidas entre los 15 y 29 años, residentes en la Comunidad de Castilla y León.

En una primera fase, consta de un *estudio cualitativo,* dirigido a captar y comprender las experiencias e imágenes que los jóvenes tienen respecto de su propia situación y entorno. Para ello se realizan 6 *grupos de discusión,* integrados por 6 u 8 jóvenes cada uno. El análisis posterior del *discurso* proporciona a los autores una valiosa información para emprender la fase siguiente de la investigación: una *encuesta* a 1.500 jóvenes, con el objetivo de conocer la dimensión estadística de cada uno de los temas del estudio.

La parte *cualitativa* comprende:

a) Un *análisis del discurso,* que permitió conocer el marco lingüístico de los jóvenes. Este análisis resultó de gran interés para garantizar el correcto funcionamiento del *cuestionario.*
b) Un *análisis de contenido,* que contribuyó a la interpretación de los resultados estadísticos de la *encuesta.* Esta información fue muy útil para la explicación de los aspectos subjetivos (actitudes, opiniones) y, en general, para descubrir el sentido que las personas atribuyen a sus conductas.

- *Torres, M.; Alvira F.; Blanco, F. y Sandí, M. (1994):* Relaciones Padres/Hijos. *Madrid, Ministerio de Asuntos Sociales.*

Una investigación interdisciplinar, en la que se materializa la *triangulación metodológica (metodología cualitativa y cuantitativa)* y la *triangulación de investigadores.* En ella convergen profesionales de diferentes ámbitos de conocimiento (sociología, psicología, estadística) para alcanzar una visión integradora de tipos distintos de relaciones entre padres e hijos en la España actual.

El estudio parte de las teorías de los modelos de socialización, configuradas en torno a dos variables fundamentales: el control y el apoyo paterno.

En una primera fase se efectúan 17 *entrevistas en profundidad* a los tres actores sociales (padres, madres e hijos), distribuidas en función de dos variables: habitat y clase social.

Las *entrevistas* se llevan a cabo en Madrid (6 urbanas), Ciudad Real (5 rurales) y Santander (3 urbanas y 3 rurales), durante el mes de octubre de 1991. La duración media de las *entrevistas* es de una hora y quince minutos.

Las *entrevistas* se realizan, en su mayoría, en los domicilios de los entrevistados. La información que de ellas se obtiene permite constatar la existencia de tres tipos básicos de relaciones paterno-filiales.

En una segunda fase se realiza una *encuesta* a 1.200 padres (70% madres y 30% padres) y a 600 hijos. El objetivo principal es analizar la distribución global de los modelos de relaciones obtenidos en la fase anterior. El trabajo de campo se efectúa durante el mes de noviembre de 1991, siguiendo el procedimiento de *muestreo de rutas aleatorias.* Los 1.200 hogares que constituyen la *muestra general* se subdividen en dos partes: la mitad en hogares con hijos entre 10 y 16 años de edad; la otra mitad en familias con hijos menores de 10 años.

El *análisis multivariable* aplicado *(análisis de componentes principales* en las respuestas a una de las preguntas del cuestionario) permitió conocer cuál es el modelo de relación padres/hijos más extendido en la sociedad española. También se obtiene los perfiles de las familias que más inciden en cada uno de los tres modelos de relaciones: el modelo autoritario, el inductivo de apoyo y el modelo *laissez-faire.*

2.2. El método científico

El objetivo de toda ciencia lo constituye el alcanzar un conocimiento que sea *científico*. Pero, ¿qué se entiende por *científico?* Para comprender su significado, resulta conveniente visualizar diferentes concepciones del *método científico.*

Uno de los primeros filósofos que delimitaron el significado del *método* en la ciencia fue Bacon. Este filósofo de la ciencia –coetáneo de Locke, Hume y Hunt– propugnó, a principios del siglo XVII, el *empirismo* como la característica más distintiva de la investigación científica. Desestima la *lógica* como fuente de verdad *(método deductivo)* y antepone, en cambio, la *experiencia,* la *observación,* como paso previo a la generalización y formación de teorías *(método inductivo).*

En 1934, las dos alternativas a la formación del conocimiento científico *(racionalismo-deducción* y *empirismo-inducción)* encuentran una síntesis en el *método hipotético-deductivo.* Concretamente, la síntesis *falsacionista* de Popper.

Este filósofo rechaza la concepción de la ciencia como saber basado en la *acumulación inductiva.* Por el contrario, en su célebre obra *La lógica de la investigación científica* (que data de 1934) defiende que "la base empírica de la ciencia objetiva no tiene nada de absoluta" (Popper, 1977: 106). Lo que confiere carácter de *cientificidad* es la solidez o fortaleza *hipotético-deductiva* de la *teoría,* su capacidad de ser rebatida por la *experiencia.*

La clave del *método científico* se halla, por tanto, en la *falsabilidad,* en "el hecho de que todas las pruebas de una teoría sean otras tantas tentativas de refutar las predicciones que se desprenden de las mismas" (Popper, 1967: 369-370).

Popper propone la prueba del *ensayo* y del *error,* de la conjetura y la refutación. Las *teorías científicas* se convierten, a partir de su aportación, en "conjeturas audaces que serán abandonadas tan sólo si entran en conflicto con posteriores observaciones" (Popper, 1967: 58).

Pero no cabe la *refutación* de una *teoría* en sentido estricto. Los hallazgos científicos siempre pueden cuestionarse. Razón por la cual nunca puede afirmarse que una *teoría* sea verdadera; sólo que es superior o inferior a otras. Depende de si supera, o no, las pruebas de comprobación que falsearon otras *teorías.*

Estos planteamientos revolucionarios de Popper fueron profundizados por algunos de sus discípulos. Principalmente, por Lakatos y Kuhn, quienes cuestionaron tanto el *verificacionismo* como el *falsacionismo empírico.*

Lakatos propone el *falsacionismo refinado* o *sofisticado,* que se ajusta más al proceder real de la actividad científica. Para cada *teoría científica* se trata ahora de "es-

pecificar los hechos que la confirman y la probabilidad de la teoría a la luz de estos hechos" (Lakatos, 1975: 208).

La verdad deja de ser absoluta. Se convierte en meramente probable, aunque el conocimiento continúa siendo comprobable. De este modo, el cálculo de *probabilidades* se convierte en criterio de demarcación, y el análisis estadístico en su herramienta básica (Beltrán, 1988).

La aceptación de una *teoría* depende –a partir de Lakatos– de si ésta posee, respecto a otras *teorías* rivales, o que la precedan, un contenido empírico excedente probado, que facilite el descubrimiento de hechos nuevos. Estas nuevas *teorías,* más "potentes", reemplazan a las anteriores. De esta forma progresa el *conocimiento científico:* mediante desplazamientos progresivos de *teorías científicas,* partiendo de un *programa de investigación.*

Los *programas de investigación* se conceptúan como progresistas o "degeneradores". Depende de si conducen, o no, al descubrimiento de hechos nuevos. Estos *programas* constan de:

a) Un *núcleo central* (o supuestos básicos), que no es falsable.
b) Unas *hipótesis* auxiliares.
c) Unos *enunciados* observacionales.

Su adecuación varía en función de si los *programas* resuelven, o no, los problemas empíricos "significativos", así como los problemas conceptuales u otros que generen.

En consecuencia, para Lakatos (1974: 209), "las *revoluciones científicas* consisten en que un *programa de investigación* reemplaza a otro (superándolo de modo progresivo)". La vieja *teoría* únicamente se abandona y reemplaza por otra nueva cuando:

a) La *teoría* se muestra ineficaz para dar una explicación significativa a los problemas que acaezcan.
b) Aparece un nuevo esquema conceptual alternativo que sí proporciona una explicación veraz al problema en concreto.

En caso de no haberse construido un nuevo esquema conceptual, o modificado el anterior, la vieja *teoría* sigue vigente. Por esta razón, al progreso científico le interesa la existencia de *programas de investigación alternativos:* "cuanto antes empiece la competencia, mejor para el progreso" (Lakatos, 1975: 267).

Kuhn, en cambio, propone el *relativismo sociohistórico* en la concepción de la ciencia. Nunca puede tenerse plena seguridad de que un determinado conocimiento sea *científico.*

Para Kuhn la ciencia avanza no por *verificación* ni por *falsación* de las *teorías* existentes; sino por *sustitución* del modelo explicativo o *paradigma* vigente por otro posterior.

La aceptación del nuevo *paradigma* dependerá, esencialmente, de la existencia de consenso (o acuerdo) en los científicos, y de la inexistencia de evidencia que lo contradiga. Esto acontece "cuando el fracaso persistente para obtener la solución de un problema importante haya producido una *crisis*. E incluso entonces, solamente se produce después de que el sentimiento de *crisis* haya producido un candidato alternativo a *paradigma"* (Kuhn, 1978a: 225).

De ahí la caracterización de Kuhn de la *actividad científica* como búsqueda de soluciones sustitutivas a los *paradigmas en crisis;* y su concepción de las *revoluciones científicas* como las responsables del avance o progreso del *conocimiento científico*.

En sus razonamientos se observan semejanzas con el otro alumno destacado de Popper (Lakatos). Los *paradigmas* actúan, al igual que los *programas de investigación,* como marco delimitador de cómo hay que proceder en el proceso de investigación. Asimismo, su sustitución exige la aparición de un nuevo *paradigma,* o la modificación de los anteriores. Este *nuevo paradigma* ha de mostrar ser más significativo, que los anteriores, en la explicación de problemas concretos.

En 1962 Kuhn publica su obra *La estructura de las revoluciones científicas,* que impulsó el cambio de orientación de la sociología a partir de los años sesenta (como ya se vio en el Capítulo 1). En esta obra Kuhn acuña el término *paradigma;* si bien, en 1969, coincidiendo con una edición en japonés de la obra citada, es cuando especifica qué entiende por *paradigma:*

> "Por una parte, pasa por la completa constelación de creencias, valores, técnicas, y así sucesivamente, compartidas por los miembros de una comunidad dada. Por la otra, denota una especie de elemento en tal constelación, las soluciones-enigmas concretas las que, empleadas como modelos o ejemplos, pueden reemplazar a reglas explícitas como base para la solución de los enigmas restantes de la *ciencia normal"* (Kuhn, 1978a: 269).

Por *ciencia normal* Kuhn (1978a: 33) comprende cualquier

> "Investigación basada firmemente en una o más realizaciones científicas pasadas, realizaciones que alguna comunidad científica particular reconoce, durante cierto tiempo, como fundamento para su práctica posterior. En la actualidad esas realizaciones son relatadas, aunque raramente en su forma original, por los libros de texto científicos, tanto elementales como avanzados".

Feyerabend va más allá. Parte de que las *teorías* no pueden ser probadas ni refutadas de forma concluyente. Se muestra contrario al *método,* al actuar conforme a unas reglas metodológicas fijas. Defiende, por el contrario, las *interpretaciones naturales* y el *conocimiento de sentido común.* Esto le lleva a propugnar la *contrainducción* (o la formación de *hipótesis* inconsistentes), y a encuadrarse dentro del *anarquismo metodológico:* "todas las metodologías tienen sus limitaciones y la única regla que queda en pie es la de que *todo vale"* (Feyerabend, 1974: 296).

Derek Phillips comparte estos mismos planteamientos. Discrepa de los supuestos metodológicos de la *ciencia positiva* de una forma radical. Propone el abandono del *método* y de la comunidad científica que lo practica:

> "Intentemos mirar el mundo con nuestros propios ojos, y no a través de instrumentos científicos que con frecuencia nublan nuestra visión" (Phillips, 1973: 114).

Desde posturas no radicales, también se transmite el cuestionamiento de la concepción tradicional del *método científico,* vinculado a la *lógica positivista.* Como apunta Alvira (1992: 59-60):

> "No existe hoy un solo sociólogo que acepte el esquema del *método científico –método hipotético-deductivo–* plenamente y sin dudas. No coincide ni con lo que hacen los científicos en las ciencias físicas o naturales, ni mucho menos con lo que hacemos los sociólogos."

En el pasado, el énfasis se ha puesto, sobre todo, en la *verificación* y *justificación* de *teorías,* y no en la *inducción,* en la *generación de nuevas teorías* a partir de *datos empíricos.* Ello ha contribuido a la carencia de "nuevos descubrimientos teóricos", con el reiterado recurso a los clásicos.

No obstante, el *método científico* se presenta hoy en día especialmente útil, no como ideal al que la práctica investigadora debiera aproximarse, sino como "referencia y patrón sobre el cual analizar los problemas y fases de la investigación en sociología" (Alvira, 1992: 62).

2.2.1. Integrantes esenciales del método científico

Sin duda, uno de los esquemas más conocidos en la comunidad científica formada por los sociólogos, en el que se representa gráficamente el proceso de investigación científica, sea el que Wallace publicara en 1971. En él se describe un proceso circular continuado, en el que quedan conectados cuatro *componentes* fundamentales del *conocimiento (teorías, hipótesis, observaciones* y *generalizaciones empíricas)* mediante cuatro *procesos cognitivos* diferentes *(deducción, operacionalización, interpretación e inducción).*

A partir de este famoso esquema han ido apareciendo otros posteriores –como el de Bryman (1988) o el de Marshall y Rossman (1989)–, que igualmente describen el *proceso de conocimiento científico* como un *proceso circular.* En él, la *teoría* y la *investigación* se condicionan mutuamente.

De acuerdo con estos esquemas (ver Figura 2.2), el proceso de investigación comienza a partir de una *teoría,* que tratará de comprobarse o verificarse. De la *teoría* se extraen (por medio de la *deducción lógica)* unas *hipótesis* concretas. Estas tendrán que ser operacionalizadas, si se pretende su *contrastación empírica.*

A continuación, se procede a la *observación* o recogida de *datos empíricos.* Estos datos serán procesados para su posterior *análisis estadístico* o *semiológico,* dependiendo de la naturaleza de la información obtenida y de la metodología escogida en el *diseño de la investigación.*

Una vez concluida esta fase, se procede a la *interpretación* de la información extraída en la investigación. De ella se deducen unas *generalizaciones empíricas* que, se-

guidamente, serán contrastadas con las *hipótesis* originales de la investigación. Si los datos empíricos contradicen las *hipótesis* del estudio, se procederá a su rechazo y a la consiguiente no aceptación de la *teoría* de la que se extrajeron las *hipótesis*. En caso contrario, las *hipótesis* serán aceptadas; pero ello no implicará la *verificación teórica,* sino sólo su confirmación –siguiendo el criterio *falsacionista* de Popper (1967)–.

De esta forma, las *generalizaciones empíricas* revierten, mediante el *método inductivo*, en la *teoría,* que se verá confirmada, rechazada o ligeramente modificada.

Cuando la *teoría* no resulta completamente apoyada por los datos empíricos recabados, el *proceso* comienza de nuevo. Ahora se tratará de comprobar (empíricamente) las modificaciones introducidas, a partir de los datos, en la *teoría* inicial de la investigación. De este modo, el proceso de adquisición del *conocimiento científico* se configura como un *feed-back* continuo. En él, la teoría y los datos empíricos se condicionan mutuamente mediante dos operaciones metodológicas esenciales:

a) La *deducción:* de la teoría a los datos, de lo general a lo concreto.
b) La *inducción:* de los datos a la teoría, de los casos particulares a los principios generales.

Esta concepción genérica del proceso de adquisición del *conocimiento científico* se resume en la Figura 2.2, que representa una adaptación del esquema original de Wallace (1971).

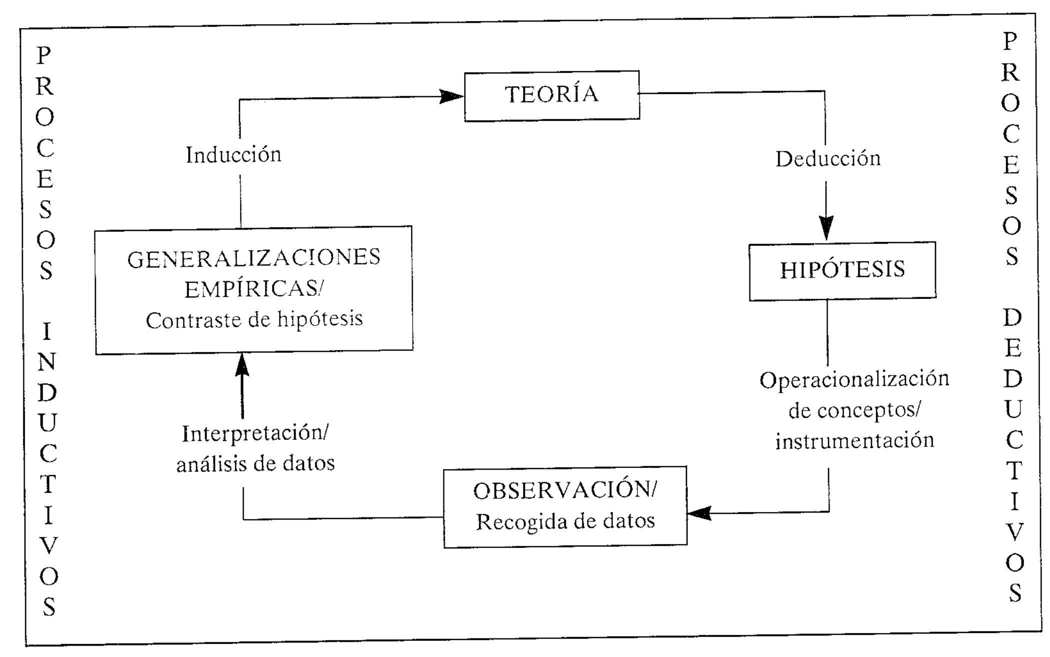


Figura 2.2. El proceso de conocimiento científico.

En suma, los integrantes esenciales del *método científico* serían, en conformidad con el esquema representado, los siguientes:

a) Teoría.
b) Hipótesis.
c) Observación.
d) Generalizaciones empíricas.

• *Teoría*

La *teoría* debería considerarse como "conglomerados integrados de proposiciones y visiones del mundo" (Layder, 1993: 15). Merton (1983: 56) la define como "grupo de proposiciones lógicamente interconectadas, de las que pueden deducirse uniformidades empíricas". Sus componentes básicos son:

a) *Conceptos:* "un símbolo lingüístico que categoriza a los fenómenos" (Phillips, 1985: 77). No constituyen *teorías,* sino "*categorías* para la organización de ideas y observaciones" (Bulmer, 1992: 43). Por lo que "sirven para clasificar y medir operativamente los fenómenos que se observan" (Blalock y Blalock, 1968: 7).
b) *Proposiciones,* que indican cómo y en qué condiciones generales se cree que se hallan relacionados los fenómenos definidos conceptualmente.
c) *Axiomas* o *leyes,* que permiten explicar las regularidades o conexiones esenciales existentes en los fenómenos sociales. (En el apartado C, dedicado a las *hipótesis,* se ilustra cada uno de estos componentes básicos de la *teoría* y su transformación en hipótesis.)

En el *proceso de conocimiento científico,* las *teorías* cumplen unas *funciones básicas* que Wallace (1971) resume en dos fundamentales:

a) *Explicar generalizaciones empíricas* ya conocidas, subsumiendo estas afirmaciones en *conceptos* y *proposiciones* apropiadas.
b) *Predecir generalizaciones empíricas* todavía desconocidas, guiando el futuro de la *ciencia.*

A estas funciones Lazarsfeld (1973a) añade la función de *interpretación:* la *teoría* permite relacionar unos *hechos empíricos* con otros *hipotéticos* o ya verificados. Además, la *teoría* resulta imprescindible en las fases preliminares de la *investigación,* a la hora de:

a) Decidir los *esquemas clasificatorios* a utilizar.
b) Plantear *conceptos teóricos* que orienten el análisis.
c) Formular *problemas de investigación* que tengan una determinada relevancia social.
d) Concretar *ideas generales* sobre cómo se producen los cambios sociales.
e) Formular *hipótesis:* realizar previsiones a partir de descubrimientos aún no verificados.

De forma similar, Merton (1967, 1983) resume en cinco las *funciones principales de la teoría* en la *investigación social:*

a) *Formular* hipótesis.
b) *Encauzar* el problema de investigación.
c) *Analizar* los conceptos sociológicos.
d) *Sugerir* interpretaciones *post-factum.*
e) *Generalizar* los resultados de la investigación.

Merton (1967) distingue dos *tipos de teorías sociológicas:* las "grandes teorías" y las "teorías de rango medio".

a) Las *grandes teorías, teorías generales o globales,* ofrecen explicaciones bastante abstractas (especulativas y poco apoyadas en la evidencia empírica) de aspectos generales de la vida social. De ahí la dificultad de su materialización global en una *investigación empírica* concreta.
No obstante, algunos aspectos específicos de estas *teorías* pueden servir de base a la *investigación* de dos formas principales:

1) Guiando el proceso de investigación.
2) Proporcionando una explicación de los hallazgos.

EJEMPLO DE "GRANDES TEORÍAS"

Merton (1967) ejemplifica esta variedad teórica con dos teorías fundamentales:

a) La *teoría marxista del desarrollo social,* porque proporciona una explicación genérica del desarrollo histórico de la sociedad moderna.
b) La *teoría del sistema social* de Parsons (1951). Esta vincula aspectos psicológicos y sociales de la conducta humana, en un intento de explicar las características definitorias de la sociedad occidental.

b) Las *teorías de rango medio* se limitan, en cambio, a aspectos concretos de la vida social, a relaciones entre variables que puedan medirse empíricamente. Por esta razón su materialización resulta más habitual en la *investigación social.*

De este tipo de *teorías* se derivan *hipótesis* que pueden ser empíricamente comprobadas.

EJEMPLO DE "TEORÍAS DE RANGO MEDIO"

Merton (1967) cita tres ejemplos clásicos de *teorías de rango medio*: *El suicidio,* de Durkheim; *La ética protestante y el espíritu del capitalismo,* de Max Weber; y, *El 18 de brumario de Luis Bonaparte,* de Marx.

A) El suicidio, de Durkheim

En esta investigación, Durkheim parte de la *observación* de que católicos de diferentes poblaciones presentan una tasa de suicidio inferior a la registrada en la población protestante. Lo mismo sucede entre la población rural, los casados, los que tienen hijos, y las mujeres, entre otros. En la búsqueda de una explicación a estas regularidades (que pueden expresarse como *generalizaciones empíricas),* se preguntó qué tenían en común las personas más propensas al suicidio.

De sus análisis resultó que todas aquellas *observaciones* particulares derivaban de un principio más genérico: las personas que se hallan menos integradas en la sociedad están más predispuestas al suicidio. De esta forma, las *generalizaciones empíricas* se convirtieron en significativas, o teóricamente relevantes, porque se pudieron relacionar o derivar de otra proposición o supuesto más abstracto: la cohesión social.

La probabilidad de que esta *inducción,* efectuada por Durkheim, sea correcta se debe a que se analizaron diversas variables, obteniéndose el mismo resultado.

Los supuestos que permitieron a Durkheim inferir, por ejemplo, la existencia de relación entre las variables tasa de suicidio y religión fueron –de acuerdo con Merton (1967: 151)– los siguientes:

a) La cohesión social proporciona apoyo psíquico a los miembros del grupo que están sujetos a ansiedades y tensiones fuertes.
b) La tasa de suicidio es función de ansiedades y tensiones no resueltas.
c) Los católicos presentan una mayor cohesión social que los protestantes.
d) En consecuencia, debería esperarse una tasa de suicidio inferior en la población católica que en la protestante.

B) La ética protestante y el espíritu del capitalismo, de Max Weber

Esta obra fue también calificada por Merton (1967: 63) como "un ejemplo fundamental de teorización (empírica) en el rango medio", por tres razones principales:

a) Trata de un problema concreto, en una determinada época histórica, aunque con implicaciones teóricas en otras sociedades y momentos históricos.
b) En ella se aplica una teoría específica sobre cómo la práctica religiosa y la actividad económica se hallan relacionadas.
c) Contribuye, al igual que la investigación de Durkheim, a una teoría más general sobre la interdependencia entre las instituciones sociales. Sin ella la teoría (de *rango medio)* no quedaría fortalecida.

C) El 18 de brumario de Luis Bonaparte, de Marx

A estos dos ejemplos clásicos de *teorías de rango medio,* Merton añade la obra de Marx titulada *El 18 de brumario de Luis Bonaparte.* Pero, a diferencia de las dos publicaciones anteriores, ésta no es razonada por el autor.

En conclusión, Merton (1967, 1983) diferencia *dos tipos de teorías* necesarias en el proceso investigador. De las *teorías de rango medio* se derivan *hipótesis* que pueden ser empíricamente comprobadas. Pero, la conclusión del proceso investigador necesitará del desarrollo de una *teoría más general o global,* que afiance los resultados de la *teoría de rango medio.*

• *Hipótesis*

A partir de una determinada *teoría* se extrae, para su *contrastación empírica,* una o varias *hipótesis* concretas. Generalmente, se recomienda extraer más de una *hipótesis,* con la finalidad de poder así abarcar los diferentes aspectos que configuran la *teoría.* Como afirman Brewer y Hunter (1989: 37), "deducir muchas hipótesis diferentes proporciona una comprobación más rigurosa del poder predictivo absoluto de la teoría, que derivar sólo una o unas cuantas hipótesis similares".

Las *hipótesis* representan *predicciones* o respuestas probables a los interrogantes que el investigador formula, ante un conocimiento previo, para su *contrastación empírica.* Popper (1977) se refiere a ellas como "predicciones deducidas de la teoría a contrastar, predicciones que han de ser fácilmente contrastables o aplicables". Suelen formularse como *proposiciones* en las que se afirma la existencia, o no, de relación esperada entre al menos dos variables en una situación determinada.

Por medio de un proceso *deductivo,* los *conceptos* y las *proposiciones* que conforman una *teoría* se concretan en *variables* e *indicadores,* que posibiliten la *contrastación empírica* de la *teoría.* Esta proceso se denomina *operacionalización de conceptos teóricos.* Consiste –siguiendo a Lazarsfeld (1973)– en:

a) Descomponer el *concepto* en las diferentes *dimensiones* o aspectos que incluye.
b) Seleccionar para cada una de dichas dimensiones una serie de *indicadores.*

c) Sintetizar los *indicadores,* mediante la elaboración de *índices* que resuman una, varias, o todas las dimensiones del concepto operacionalizado numéricamente (esto se trata con mayor detalle en el Capítulo 4).

Además de las *variables,* en las *hipótesis* se hace también referencia, aunque genérica, a las *unidades de observación.* Las unidades pueden ser individuos o agregados. En la investigación social, la *unidad de observación* más habitual es el individuo. De él se extrae información sobre sus características personales y de su entorno sociocultural.

Debido al elevado coste (en tiempo y dinero) que supone una recogida de información exhaustiva, el recurso a una *muestra* es la norma. Claro que en todas las *investigaciones* no se siguen los mismos criterios de *selección muestral.* Depende, fundamentalmente, de cuáles sean los objetivos del estudio y de los recursos económicos y humanos disponibles (como se verá en el Capítulo 5).

EJEMPLO DE FORMULACIÓN DE HIPÓTESIS

Como ilustración de lo expuesto, situémonos en una investigación sobre causalidad de la delincuencia juvenil. Una de las *teorías* que mayor apoyo empírico han recibido (Hindelang, 1973; Johnson, 1979) es la *teoría del control social* de Hirschi (1969). Algunos de los *axiomas* (o aseveraciones esenciales) en los que se fundamenta esta *teoría* son los siguientes:

a) Los *controles sociales* son los que mantienen al joven en la conformidad.
b) La *interiorización de la norma* constituye el determinante de la conducta delictiva.
c) Los *lazos* o *frenos* que mantienen al individuo en la conformidad son:

- *Afecto* a personas convencionales (padres, profesores, amigos).
- *Compromiso* con líneas convencionales de acción.
- *Implicación* en actividades convencionales.
- *Creencia* en la legitimidad.

De estos comienzos axiomáticos, podrían deducirse unas *proposiciones,* a partir de las cuales se derivan *hipótesis* como las siguientes:

a) Cuanto mayor respeto y admiración siente el joven hacia las personas con las que se identifica, menor es la probabilidad de que se emprenda en conductas delictivas.
b) Cuando el joven se libera de las ataduras que le vinculan a la sociedad convencional (familia, escuela, comunidad), está más predispuesto a transgredir la norma legal.

c) Cuanto más ocupado esté el joven en la realización de actividades convencionales, menor tiempo encontrará para cometer actos delictivos.
d) La ausencia de creencias que prohíban la realización de conductas delictivas, lleva al joven desvinculado a la delincuencia.

Las *hipótesis* contienen *conceptos* que necesitarán medirse para su comprobación. Este proceso se conoce como *operacionalización* o traducción de conceptos generales en *indicadores* concretos. En este ejemplo, podría decidirse *operacionalizar los conceptos:*

a) *Delincuencia juvenil,* como todo tipo de conducta que infringe la norma legal llevada a cabo por personas menores de 18 años de edad.
b) *Falta de control paterno,* como grado al que el joven percibe que sus padres le rechazan o no apoyan.

Téngase presente que cualquier *operacionalización* que se efectúe de un *concepto teórico* es difícil que cubra todas las *dimensiones* que representan al *concepto* en cuestión. Se tratará siempre de una aproximación susceptible de ser rebatida. En el *ejemplo* anterior, la manera como se ha operacionalizado el *concepto* de *delincuencia juvenil* (transgresión de la norma legal) no sería compartida por los seguidores del *paradigma interaccionista* (Teoría de etiquetamiento). Para estos autores, la reacción negativa de la sociedad es la que define la conducta delictiva, no la transgresión de la norma. Una definición más amplia de la conducta delictiva sería la proporcionada por autores como Gibbons y Jones (1971), que la definen como aquel tipo de conducta que:

a) Supuestamente viola normas de conducta generales, desde el punto de vista cultural.
b) Origina fuertes reacciones sociales.
c) Se manifiesta en actividades de control social formal hacia dicha conducta en manos de la policía y similares.
d) A menudo lleva a la *desviación secundaria* (al rol de desviado).

(En el artículo de Alvira (1975), "La teoría del etiquetamiento y el concepto de conducta desviada", *Revista de Opinión Pública,* nº 39; puede encontrarse resumida la polémica en torno al concepto de conducta desviada).

• *Observación*

Una vez que se ha seleccionado el ámbito de la realidad social a analizar, se procede a la *observación* o *recogida de información.* El investigador recoge datos relacionados con los *indicadores empíricos* seleccionados para su posterior análisis, en conformidad con la *metodología* elegida en el *diseño de la investigación.* Bien mediante *cuestionario autoadministrado, entrevista personal,* utilización de *datos disponibles* (in-

formes policiales, judiciales, escolares) o cualquier otra técnica de recogida de datos que el investigador estime necesaria; o bien una conjunción de diversas técnicas, como ya se vio en la sección 2.1.2.

• *Generalizaciones empíricas*

El círculo del proceso de investigación científica se cierra cuando los resultados de la investigación revierten, mediante la operación metodológica de la *inducción*, en la *teoría* inicial. La información que se obtuvo en la fase de *observación* se interpreta, ahora, en busca de *generalizaciones empíricas*. Por *generalizaciones empíricas* Merton (1983: 171) entiende

> "Cualquier proposición aislada que resuma uniformidades observadas de relaciones entre dos o más variables".

A continuación, se procede a contrastar las *generalizaciones empíricas* con las *hipótesis* de la *investigación,* mediante un proceso inferencial conocido como *contraste de hipótesis*.

Cuando los datos empíricos corroboran las *hipótesis teóricas,* ello implica la confirmación de la teoría. En caso contrario, la *teoría* original queda debilitada por los hallazgos empíricos de la investigación, lo que conlleva su modificación o rechazo. Como apunta De Vaus (1990: 19-20):

> "Rara vez la teoría inicial es completamente apoyada por la investigación. Esto es bueno porque nos hace pensar, modificar o desarrollar la teoría inicial y es lo que nos lleva al progreso. Como tal la teoría modificada necesitará ser comprobada rigurosamente."

De esta forma, el proceso de investigación comienza de nuevo. Pero, téngase presente que "cualquiera que sea el método utilizado en la contrastación de hipótesis, los resultados nunca son *ciertos,* sino aproximaciones en términos de *probabilidad*" (Goode y Hatt, 1952: 87). Por esta razón, para cada *teoría* se tratará de "especificar los hechos que la confirman y la probabilidad de la teoría a la luz de estos hechos" (Lakatos, 1975: 208).

Merton (1967, 1983) resume en tres las formas principales en que la *investigación empírica* ayuda a la *formulación de la teoría:*

a) Lo que denomina *serendipity pattern:* aquella situación en la cual hallazgos de investigación no anticipados (que no son consistentes con la *teoría* prevaleciente u otros hechos establecidos) estimulan el desarrollo de una *nueva teoría* o la modificación de la *teoría* existente.

b) La *observación* repetida de hechos que anteriormente se habían negado presiona hacia la modificación o ampliación de la *teoría*.

c) El proceso de *investigación* fuerza al investigador a establecer *definiciones operacionales* de *conceptos*. Esto ayudará a la clarificación de la *teoría* o *marco teórico*.

Merton es uno de los autores que más insiste en la necesidad de que los investigadores desarrollen una orientación teórica: realicen investigaciones en el contexto de un *marco teórico,* comprobando *teorías* y modificándolas a la luz de los resultados de la *investigación*.

Sin duda, la *teoría* y la *investigación* han de interrelacionarse. Toda *teoría* ha de basarse en *hallazgos empíricos* para no quedar en una mera especulación ideológica. De igual manera, toda *investigación* debe partir de una *teoría* para adquirir una mayor relevancia. Pero esta necesidad mutua de la teoría y la investigación no siempre se ha reconocido. Si bien, en los últimos años su reconocimiento ha aumentado.

En 1981 Wells y Picou publicaron un *análisis de contenido* de una *muestra* de 707 artículos de los 2.619 que fueron publicados en la revista *American Sociological Review* entre 1936 y 1978. El objetivo del análisis era averiguar cómo la *teoría* ha figurado en la *investigación sociológica*. Los autores del estudio obtuvieron que la presencia de la *teoría* en el proceso de *investigación* ha aumentado progresivamente desde 1936.

Si entre 1936 y 1949 sólo un 34% de los artículos publicados en la revista incluían alguna utilización de la *teoría,* entre 1965 y 1978 supuso un 66% de los artículos. Asimismo, hallaron un aumento espectacular en el modelo de *comprobación de hipótesis* derivadas teóricamente (lo que constituye el *modelo tradicional* de indagación científica). Si durante 1936-1949 este modelo se utilizó sólo en el 16% de los artículos, la proporción aumenta hasta el 54% en los años 1965-1978.

Por último, conviene puntualizar –y sobre ello se volverá en el capítulo siguiente– que no todas las *investigaciones empíricas* que se llevan a cabo cumplen cada una de las fases o etapas enunciadas (en esta sección), y de forma lineal. Depende del estudio (si se fundamenta en *datos secundarios* o *primarios)* y, en especial, de la *metodología* seguida *(cuantitativa o cualitativa)*.

Ambas *metodologías* comparten unas fases similares en su desarrollo básico: selección y definición del *problema* a partir de una *teoría* o *investigación* previa; formulación de *hipótesis;* elección del *diseño de investigación;* recogida de *datos; análisis* de la información; *validación;* e *informe*.

No obstante, la *metodología cualitativa* se caracteriza por una redefinición constante, tanto del *objeto* de estudio como de las *hipótesis* de investigación, hasta lograr su adecuación. De ahí que las diferentes fases del *proceso de conocimiento científico* no acontezcan de manera gradual, sino interactivamente. El análisis de los datos empíricos puede implicar la modificación de las *hipótesis* de partida, con la consiguiente recogida de información y, así, sucesivamente.

Lo mismo acontece en la *investigación cuantitativa*. En ella tampoco se sigue siempre, y de forma rígida, el esquema señalado.

En suma, en la práctica de la *investigación social* no siempre se asiste a una sucesión ordenada de fases. Se trata más bien de un *feed-back* continuo.

2.2.2. La cientificidad de la sociología

Aunque existan autores, como Giner (1979: 20), que responde categóricamente "sí", "la sociología es una ciencia porque cumple con el sentido básico de lo que el vocablo latino *scientia* expresa: saber *objetivo* y *racional* de la realidad"; también, porque "reúne las condiciones mínimas de *empirismo, teoría, apertura, neutralidad ética* y *crítica* que de toda ciencia se exige"; lo cierto es que en ésta, como en otras cuestiones, no existe unanimidad entre los distintos autores. Especialmente, en lo concerniente a la *objetividad* o *neutralidad ética*. Este aspecto, unido a los *problemas de interacción* y de *medición*, ha contribuido a agudizar la polémica en torno a la *cientificidad* de las ciencias sociales, basada en el referente de las ciencias naturales. Incluso algunos autores, como Feyerabend (1974) o Phillips (1973), llegan a negar la *objetividad* de la ciencia social y su posibilidad como *conocimiento científico*.

Lo que distingue a la *ciencia* respecto de otra forma de conocimiento no es su objeto, sino su *método:* su capacidad de explicar la realidad conforme a unos principios lógico-empíricos, aceptados por la generalidad de la comunidad científica de referencia.

Entre las características que ha de reunir el *conocimiento* para adquirir la condición de *científico*, Gibson (1982) señala las siguientes: *abstracción, generalidad, observación experimental* o *evidencia empírica, neutralismo ético* y *objetividad*.

De acuerdo con Bottomore (1978), el *conocimiento* sociológico puede considerarse *científico* por tres razones fundamentales:

a) Se ocupa de hechos (no de juicios de valor).
b) Aporta pruebas empíricas en apoyo de las afirmaciones.
c) Es objetivo.

Estos criterios de *cientificidad* concuerdan con el *paradigma* que ha prevalecido en la sociología durante años *(paradigma cuantitativo)*, no con el *paradigma cualitativo*.

Los teóricos críticos, por ejemplo, admiten que la sociología sea una *ciencia*, aunque niegan el *avalorismo*. Tratan de llegar a la verdad, pero mediante el reconocimiento de los *valores*. En su opinión, la *ciencia* debe descubrir lo que hay detrás de los hechos, operando mediante la autorreflexión. Por tanto, estos autores admiten la *objetividad*, pero como medio para alcanzar el verdadero objeto. "La objetividad de la investigación empírica no es, por lo general, sino la objetividad de los métodos, no de lo investigado" (Adorno, 1973: 84).

El hecho de que el hombre sea a su vez objeto y sujeto de conocimiento es lo que dificulta la *objetividad* en las ciencias sociales. "Este objeto, a su vez, es un sujeto cons-

ciente, dotado de palabra y de múltiples simbolismos, con lo que la objetividad y sus previas condiciones se hacen tanto más difíciles y a menudo limitadas" (Piaget, 1973: 67).

Por ésta y otras razones, Bruyn (1972) prefiere obviar el tema de la *objetividad* y apunta a la existencia de "coherencia teórica, correspondencia empírica y consenso por experiencia", a modo de instancias genéricas en las que debe asentarse la adecuación del *conocimiento científico*.

En resumen, se reconoce la influencia de los *valores* en la toma de decisiones iniciales en el *proceso de investigación* (adopción de criterios fundamentales y elección del tema de la investigación). Lo que provoca la polémica es su aceptación en el contenido de los enunciados sociales científicos.

Ya Weber (1977) se pronunció a favor de los *valores* como útiles en el *proceso de investigación,* pero en cuanto instrumentos de selección del *conocimiento*. El análisis, por el contrario, debería realizarse de la manera más neutral posible. Ello debe -en opinión de Weber (1981)- llevar al científico, a diferencia del político, a explicitar sus propios *valores* para que no se confunda la realidad que describe con los ideales que profesa.

Asimismo se pronuncian otros autores (como Mannheim, 1958; Myrdal, 1970 o Gouldner, 1971), quienes admiten la *cientificidad* de la *sociología* e, igualmente, recomiendan al investigador explicitar aquellos *valores* de los que no se pueda liberar.

Los *valores* no introducen sesgos en la *investigación* cuando señalan problemas, sino cuando prejuzgan soluciones (Nagel, 1961; Kaplan, 1964). De ahí que, "la introducción de ciertos juicios morales en la investigación de la realidad social puede disminuir fácilmente la calidad de los resultados" (Giner, 1979: 23).

A modo de conclusión puede afirmarse que la *sociología* sí constituye una *disciplina científica* (puede desarrollar *teorías* verificables científicamente). Si bien, el sociólogo ha de procurar analizar la realidad social de la forma menos parcial y lo más objetiva posible, si pretende que los hallazgos que obtenga adquieran credibilidad y reconocimiento por parte de la comunidad científica.

Lecturas complementarias

Alvira, F. (1983): "Perspectiva cualitativa-perspectiva cuantitativa en la metodología sociológica", *REIS*, n.º 22, pp. 53-75.

Beltrán, M. (1988): *Ciencia y Sociedad*. Madrid, CIS.

Brewer, J. y Hunter, A. (1989): *Multimethod research. A synthesis of styles*. London, Sage.

Bryman, A. (1988): *Quantity an quality in social research*. London, Unwin and Hyman.

Cook, T. y Reichardt, Ch. (1986): *Métodos cualitativos y cuantitativos* en investigación *evaluativa*. Madrid, Morata.

Denzin, N. K. (1975): *The research act: a theoretical introduction to sociological methods*. Chicago, Aldine Publishing Company.

García Ferrando, M. (1979): *Sobre el método*. Madrid, CIS.

Merton, R. K. (1983): *Teoría y estructuras sociales*. México, FCE.

Wallace, W. (1976): *La lógica de la ciencia en la sociología*. Madrid, Alianza.

Ejercicios propuestos

1. ¿En qué consiste la *triangulación multimétodo?* Ejemplifíquelo con una investigación.
2. Señale *sesgos* que puedan surgir en una investigación *multimétodo*.
3. ¿Qué se entiende por "revolución científica"? ¿En qué difieren los planteamientos de Kuhn de sus predecesores y coetáneos?
4. ¿Qué da carácter "científico" a un conocimiento? Justifique la respuesta.
5. ¿Qué teorías son de "rango medio"? Escoja una teoría y extraiga los axiomas principales en los que se fundamenta. A partir de ellos, formule unas hipótesis e indique los conceptos que habría que operacionalizar.

SEGUNDA PARTE

LOS PRELIMINARES DE LA INVESTIGACIÓN SOCIAL

3

LA ORGANIZACIÓN DE LA INVESTIGACIÓN

Una vez definido el *método científico* y sus condicionantes, corresponde ahora pasar del proceso de investigación "ideal" a la práctica habitual. Esta se halla metodológicamente determinada por los *objetivos* y el *marco teórico* del estudio. Pero también intervienen aspectos económicos (los *recursos* a disposición del investigador) y temporales (el plazo de *tiempo* concedido para su materialización).

Todos estos aspectos marcan el desarrollo de la investigación, desde su configuración inicial en un *proyecto de investigación.* Concretamente, afectan al *diseño:* a la selección de *estrategias*, con las correspondientes *técnicas* para la recogida y el análisis de la información.

Pero, ante todo, el *proceso de investigación* no ha de concebirse como algo fijo e inmutable. Aunque el *proyecto* haga una expresa relación de las actuaciones para el cumplimiento de cada fase de la indagación, su puesta en práctica puede llevar a la modificación –a veces sustancial– del diseño original. Razón por la cual se recomienda la previsión (en el *proyecto de investigación),* de posibles incidencias en su ejercicio, así como *sesgos* previsibles que pudieran invalidar los resultados del estudio. Éste es el *desideratum.*

3.1. El proyecto de investigación

Dos conceptos parejos suelen, a veces, confundirse: el de *diseño* y el de *proyecto.* Si bien, este último constituye un concepto más amplio que el anterior.

En el ámbito de la *investigación social,* el concepto de *diseño* se restringe a la *planificación* de la *estrategia* para alcanzar los *objetivos* de la investigación. Alvira (1989: 85) lo define como un "plan global de la investigación que integra, de modo co-

herente y adecuadamente correcto, técnicas de recogida de datos a utilizar, análisis previstos y objetivos".

Igualmente, Yin (1989) lo describe como una secuencia lógica que relaciona los datos empíricos del estudio con las preguntas formuladas al inicio de la investigación y, en último término, con las conclusiones. "Se trata de un problema lógico, no logístico, y por tanto es más que un plan de trabajo" (Yin, 1989: 29).

La función esencial del *diseño* es guiar al investigador en la obtención y posterior análisis de la información. El *proyecto* se configura, en cambio, como un conjunto de elementos entre los cuales se encuentra el mismo *diseño*.

Además de englobar al *diseño*, el *proyecto* abarca otros aspectos varios: desde la idea originaria de la investigación y su concreción en objetivos específicos, hasta el detalle económico, organizativo y temporal de las diferentes fases del proceso de investigación.

El origen de la confusión entre estos dos conceptos se halla –como bien explica Alvira (1989: 17)– en el hecho de que la elección y elaboración de un *diseño de investigación* "necesita contemplar todos los extremos que un *proyecto* presenta". Razón que justifica que ambos términos con frecuencia se "confundan a efectos prácticos".

En resumen, el *proyecto de investigación* se hace en referencia a tres elementos claves:

a) Los *objetivos* de la investigación.
b) Los *recursos* o medios materiales, económicos y humanos de que disponga el investigador para hacer viable el estudio.
c) El *tiempo* concedido para su realización.

La conjunción de estos elementos marcará el alcance y desarrollo de la investigación, por su incidencia directa en el *diseño de la investigación*.

En el Cuadro 3.1 se resumen los componentes esenciales de un *proyecto de investigación*. Asimismo, en la Figura 3.1 (al final de la presente sección) se representa, de forma gráfica, la configuración del *proyecto de investigación*. A continuación se incluye el detalle de cada uno de estos integrantes.

3.1.1. La formulación del problema de investigación

> "No importa en qué problema quieras trabajar y no importa qué método eventualmente vayas a utilizar, tu trabajo empírico deberá comenzar con una consideración cuidadosa del problema de investigación" (Simon, 1978: 98).

En concreto, ha de exponerse:

a) Qué es exactamente lo que se pretende analizar: cuáles son los *objetivos generales* y *específicos* de la investigación.

b) Qué *razones* motivaron la elección de ese *problema* concreto.
c) La *justificación* de dicha elección, relacionándola con la relevancia y trascendencia de la indagación.

Esta exposición de *razones* y *justificaciones* ha de hacerse de forma convincente, en aras de lograr la aprobación del *proyecto de investigación,* y su consiguiente financiación económica (en caso de solicitarse).

Al principio, la *idea* o *problema de investigación* se presenta de forma vaga y genérica. Pero, paulatinamente, ira concretándose, conforme el investigador vaya ahondando en ella.

La *fuente originaria de la idea* puede ser muy dispar. A veces es la propia entidad o institución que financia la investigación quien propone el *problema* o *tema de estudio.* En otras ocasiones, es el azar quien suscita la *idea:* en el curso de una conversación; durante la lectura de un libro o artículo de prensa (periódico, revista); o tras la observación de un acontecimiento, de forma directa o indirecta (por medio de la televisión, la radio), entre otras posibles *fuentes originarias de ideas.*

Pero la *idea* también puede ser el resultado de una reflexión profunda, presente durante mucho tiempo en la mente del investigador, aunque hasta entonces no haya sido concretada ni operacionalizada en un *proyecto de investigación* específico.

Al principio, sólo es eso: una mera *idea.* El investigador, posteriormente, tendrá que ir precisándola y configurándola. Para ello deberá adentrarse en el campo de conocimiento determinado donde se ubique la *idea* en cuestión.

CUADRO 3.1. Componentes esenciales de un *proyecto de investigación.*

1. La *formulación del problema de investigación:*
   - Definición de objetivos.
   - Elaboración del *marco teórico* de la investigación: revisión bibliográfica y demás indagaciones exploratorias.
2. La *operacionalización del problema:*
   - Formulación de hipótesis.
   - Operacionalización de conceptos teóricos.
   - Delimitación de las unidades de análisis.
3. El *diseño de la investigación:*
   - Selección de *estrategias.*
   - Diseño de la *muestra.*
   - Elección de *técnicas* de recogida y de análisis de datos.
4. La *factibibidad de la investigación:*
   - Fuentes bibliográficas.
   - Recursos disponibles (materiales y humanos).
   - Recursos necesarios (económicos, materiales, humanos).
   - Planificación del tiempo de realización.

A tal fin se recomienda efectuar una *revisión bibliográfica* exhaustiva sobre el *tema* concreto de interés y otros similares que, aunque no se refieran al *problema* específico del estudio, sí puedan orientar en su indagación. Esta última *revisión de temas* relacionados con el *problema de investigación* adquiere una mayor relevancia cuando apenas se encuentra *bibliografía* (teórica y empírica) específica sobre dicho tema, bien debido a su carácter novedoso, o bien a su escasa repercusión en la tradición investigadora anterior.

La importancia de llevar a cabo una *revisión bibliográfica* exhaustiva es destacada por la mayoría de los autores. Incluso llega a afirmarse que "todo proyecto de investigación en las ciencias sociales debería implicar la búsqueda de indagaciones previas" (Cooper, 1984: 7).

Con la *revisión bibliográfica* se busca la familiarización con el *tema de estudio* escogido: hallando, leyendo, evaluando y sintetizando indagaciones realizadas con anterioridad. De ellas se extraerá información concerniente a:

a) La *metodología* seguida y los *resultados* que, mediante ella, se lograron: sus potencialidades y limitaciones concretas.
b) Aspectos del *problema* tratados, aquellos no abordados y los suscitados tras la conclusión de la investigación, como relevantes para una futura indagación.

El propósito de dicha *revisión bibliográfica* Dankhe (1989) lo resume en uno fundamental: obtener información para, a partir de ella, poder analizar y discernir si la *teoría* existente y la *investigación* previa sugieren una respuesta (aunque sea parcial) a las *preguntas de investigación,* o una dirección a seguir dentro del estudio. Más concretamente, la *revisión bibliográfica* contribuye a:

a) La estructuración de la *idea* originaria del *proyecto de* investigación, contextualizándola en una *perspectiva teórica y metodológica* determinada.
b) Inspirar cómo *diseñar la investigación,* a partir de la experiencia obtenida en la utilización de determinadas *estrategias de investigación*. Se trataría, por tanto, de:

    1) *Prevenir* posibles *errores* cometidos en estudios anteriores al presente.
    2) *No ser redundantes*. La nueva *investigación* debería dar un enfoque distinto al proporcionado en indagaciones anteriores. Por ejemplo, escogiendo una *estrategia de investigación* diferente, con el propósito de comprobar si con ella se alcanzan los mismos resultados.

    También, se trataría de plantear otros interrogantes de investigación, con objeto de cubrir aspectos menos observados, y no aquéllos continuamente analizados.

Además de la necesaria *revisión bibliográfica,* la estructuración de la *idea de investigación* requiere el complemento de otras *indagaciones exploratorias*. Éstas pueden materializarse en:

a) *Discutir el tema con otros investigadores* que puedan aportar *ideas* de cómo llevar a cabo la investigación y qué pasos seguir.
b) *Entrevistar a algunos de los implicados* en el *problema* a investigar, con el propósito de recabar, de ellos, información que ayude al *diseño de la investigación*.

Todas estas tareas previas en el *proyecto de investigación* (desde la *revisión bibliográfica* a las *entrevistas abiertas)* se convierten en cruciales, cuanto más vaga se muestre la *idea* originaria del estudio, y más desconocedor sea el investigador del área de conocimiento específica donde ésta se ubique.

A partir de la *revisión bibliográfica* y otras *indagaciones exploratorias* (que el investigador opte por seguir), se configurará el *marco teórico* de la investigación. Éste se basará en la integración de toda la información previa que el investigador recopile. Además, le proporcionará un *marco de referencia* al que acudir en la interpretación de los resultados que obtenga tras la conclusión de la investigación.

A veces se partirá de una *teoría;* otras veces, se decidirá imbricar el *problema* de estudio en varias *perspectivas teóricas* para, de este modo, poder proporcionar cobertura a los distintos aspectos en que el *problema* se manifiesta (máxime cuanto más complejo sea el *tema de la investigación).* Pero, en más ocasiones de las deseables, el *marco teórico* que configura el *proyecto de investigación* se fundamenta en *generalizaciones empíricas: proposiciones* que han sido comprobadas en anteriores investigaciones, si bien no han logrado cristalizar en una *perspectiva teórica* determinada.

EJEMPLO DE CONFIGURACIÓN DE UN PROBLEMA DE INVESTIGACIÓN

Como ilustración de lo expuesto hasta ahora, piénsese en una *idea* que se desee desarrollar en un *proyecto de investigación.* Cualquier hecho o acción social puede convertirse en un *problema de investigación.*

La "violencia callejera juvenil" constituye uno de los fenómenos sociales más presentes en la cotidianidad de la sociedad actual. Al menos es, junto con el problema de la corrupción, uno de los más comentados en los informativos (de radio, televisión, prensa escrita), y frecuente tema de coloquio.

El conocimiento, de manera directa o indirecta, de algún acto nuevo de violencia juvenil puede despertar el interés de algún investigador deseoso de encontrar respuestas a *interrogantes* suscitados ante dichos sucesos: ¿por qué los jóvenes actuaron de la forma que lo hicieron?; ¿qué factores inciden en la comisión de actos de violencia?; ¿incita la sociedad actual a la violencia?; ¿qué características diferencian a los jóvenes violentos de los no violentos?; ¿promueve la frecuente aparición en televisión de actos violentos conductas de imitación?

Lo primero que el investigador ha de hacer es formular el *problema de investigación,* inicialmente titulable "la violencia callejera juvenil". Para ello, deberá centrarse en *tres tareas principales:*

a) *Establecer los objetivos generales y específicos del estudio:*

— *Objetivo general:* analizar las causas de la violencia juvenil.
— *Objetivos específicos*:
- Describir las características sociodemográficas y psicosociológicas del joven violento.
- Indagar en la trayectoria vital del joven (afectiva, laboral, ocupación del tiempo).
- Analizar su ambiente sociofamiliar (familia, amigos, barrio).

b) *Delimitar el tema de estudio* (qué actos de violencia se analizarán) y las *unidades de observación* (qué tipo de jóvenes, de qué edades, de qué ámbito territorial).
c) *Elaboración del marco teórico de la investigación:* qué enfoque teórico y metodológico se dará al análisis de la violencia juvenil.

Para que esta fase inicial y esencial en el *proyecto de investigación* se efectúe de forma rigurosa, el investigador deberá precederla de una *indagación exploratoria*. Ésta se convertirá en referente del *diseño de investigación*. Concretamente:

a) Una *revisión bibliográfica* de las *perspectivas teóricas* existentes e *investigaciones empíricas* llevadas a cabo en España y en otros países (que el investigador especificará) acerca de "la violencia juvenil".
Esta tarea ardua en el pasado se ve, en el momento presente, bastante facilitada por el uso generalizado de la *búsqueda bibliográfica por ordenador mediante CD-ROM*. Este proporciona información existente en bibliotecas y centros de datos de distintos países, en un breve período de tiempo.
b) *Entrevistas abiertas a expertos* en la materia y a *protagonistas* de acciones de violencia (jóvenes implicados en actos de violencia callejera y personas que han sido víctimas de dichos actos), que constituyen la *población de estudio.*

Con la *indagación exploratoria* se busca familiarizarse con el *tema de la investigación* y con la *metodología* a seguir en su análisis. De ahí que estas primeras tareas resulten esenciales para realizar un buen *diseño de investigación*. De lo exhaustivo que se sea en esta fase previa del *proyecto de investigación* dependerá el buen desarrollo de la misma.

3.1.2. La operacionalización del problema de investigación

Después de la definición y delimitación del *objeto de estudio,* procede su concreción, su operacionalización, tanto en su *vertiente teórico-analítica* (principales *conceptos, categorías analíticas* y *variables* a analizar) como *poblacional* (qué *población* se estudiará y, en su caso, qué *diseño muestral* se efectuará).

Del *marco teórico* de la investigación se extraen, mediante un proceso deductivo, unas *hipótesis* que representen respuestas probables que el investigador avanza a las *preguntas* iniciales de la investigación. Las *hipótesis* vienen expresadas en forma de *proposiciones,*

en las que se afirma la existencia o inexistencia de asociación esperada, al menos entre dos variables (dependiente e independiente, generalmente), en una situación determinada.

Se recomienda que las *hipótesis* se encuentren relacionadas con los *objetivos de la investigación* y sean lo más concretas y precisas posibles.

EJEMPLOS DE FORMULACIÓN DE HIPÓTESIS

- Los jóvenes que en su infancia han sido víctimas de malos tratos presentan, en su juventud, actitudes más violentas que aquéllos que no los han padecido.
- El consumo de bebidas alcohólicas aumenta la probabilidad de cometer actos de violencia.

Las *hipótesis* contienen *conceptos* que determinan lo que debe analizarse. Para ello las *hipótesis* han de traducirse en *variables empíricas* o *indicadores*. A este proceso se le denomina, normalmente, *operacionalización de conceptos teóricos*. Ésta constituye una fase intermedia que vincula la *formulación teórica del problema* con la *medición de las variables* relevantes. En el Capítulo 4 se tratarán con más detalle el *proceso de medición* y de *operacionalización de conceptos*.

EJEMPLO DE OPERACIÓN DE CONCEPTOS TEÓRICOS

Se parte de la *hipótesis* siguiente:

> "Los actos de violencia juvenil callejera crecen, conforme aumenta la insatisfacción social en los jóvenes."

Esta *hipótesis* contiene *conceptos teóricos* que deberán operacionalizarse (traducirse en *variables empíricas* o *indicadores),* antes de proceder a la recogida de información.

Uno de los *conceptos* es el de "insatisfacción". Habría que seccionarlo en las diferentes *dimensiones* que comprende: insatisfacción laboral, educativa, afectiva, económica, familiar, entre otras posibles.

Para cada una de estas *dimensiones* se buscarán *indicadores* que ayuden en su concreción. Por ejemplo, la "satisfacción laboral" podrá medirse por los siguientes *indicadores:*

- Situación laboral: activo, en paro.
- Tipo de contrato: fijo, eventual, a tiempo completo, a tiempo parcial.
- Trabajo que realiza en la empresa: si corresponde o no a su formación académica y técnica.
- Remuneración que percibe.

- Relación con los compañeros.
- Relación con el jefe o personal a cuyas ordenes se halle.
- Ubicación de la empresa: entorno físico dentro y fuera de la empresa.

Además de las *variables,* en las *hipótesis* también se hace mención (aunque genérica) de las *unidades de análisis:* la *población* o *universo* objeto de estudio.

La elección de una *población* u otra puede tener implicaciones significativas, tanto para el contenido sustantivo de la investigación como para sus costes (económicos y de tiempo). Por esta razón, autores como Hakim (1994) recomiendan que la *población de estudio* se establezca en los *objetivos de la investigación,* debido a que ésta determina la información a extraer y la naturaleza de los resultados de la investigación.

Las *unidades de análisis* no sólo incluyen el *ente* individual o colectivo que se observará, sino también el *espacio* donde se ubica y el *tiempo.* En los *estudios longitudinales* el *tiempo* figura (explícitamente) como una de las *variables* en la investigación.

3.1.3. El diseño de la investigación

Una vez que se ha formulado y operacionalizado el *problema de investigación* (concreción de los *objetivos,* las *hipótesis,* las *variables* y las *unidades de análisis),* corresponde decidir el *diseño de la investigación:* especificar cómo va a realizarse la investigación.

Esta fase es clave dentro del *proyecto de investigación* y, al igual que las precedentes, vendrá determinada por los *objetivos del estudio* (exploratorios, explicativos, predictivos, evaluativos), los *recursos* (materiales, económicos, humanos) y el plazo de *tiempo* disponible para su materialización. De ahí que el investigador deba siempre tener en mente -al ocuparse del *diseño*- no sólo los *objetivos* del estudio, sino también la *factibilidad del diseño* elegido. Ha de procurar que el *diseño* sea viable, con los recursos disponibles y dentro de los límites de tiempo fijados.

Esta triple consideración de *objetivos, recursos* y *tiempo* incide, directamente, en la selección de una o varias *estrategias de investigación: uso de fuentes documentales y estadísticas, estudio de casos, encuesta, experimento.*

La *estrategia* (o *estrategias)* finalmente elegida influirá en:

a) El *diseño muestral:* la *muestra* del estudio (individuos, viviendas, entidades sociales, acontecimientos, documentos), su volumen y forma de selección.

b) Las *técnicas de recogida de información* (de documentación, observación y entrevista). Concretamente:

   1) *Revisión de fuentes de observación secundaria* (estadísticas y documentos).
   2) *Observación sistemática, participante* y *no participante.*
   3) *Entrevistas abiertas, semi* o *no estructuradas* (individuales y/o grupales).

4) *Relatos biográficos* (múltiples, paralelos, cruzados) y *documentos personales.*
5) *Cuestionario estandarizado* (sondeos en serie, tipo panel, macroencuestas).

c) Las *técnicas de análisis de datos:*

1) *Documental.*
2) *Estadístico* (univariable, bivariable, multivariable).
3) *Estructural* (análisis del discurso, etnografía, fenomenografía).
4) *Interpretacional* (construcción de teorías, descriptivo/interpretativo).
5) *De contenido* (cuantitativo y cualitativo).

Las *técnicas cuantitativas de recogida y de análisis de datos* constituyen el eje de este manual. En capítulos sucesivos se expondrá cada una de ellas. Para el conocimiento, en cambio, de las *técnicas cualitativas* remito al lector interesado al manual de Miguel S. Valles Martínez, publicado en esta misma colección.

En suma, en la selección de la *estrategia* y de las *técnicas* (de recogida y de análisis de datos) correspondientes intervienen los siguientes elementos:

a) Los *objetivos de la investigación* y si ésta se ceñirá a un momento temporal concreto *(diseños seccionales o transversales)* o si, por el contrario, incluirá diferentes períodos de tiempo *(diseños longitudinales).*
b) El *grado de conocimiento*, por el investigador, de la diversidad metodológica existente.
c) El *grado de rigor y precisión* que el investigador desee para su indagación.
d) La *factibilidad de la investigación.* Ésta se concreta, siguiendo a Pons (1993: 24), en los factores siguientes:

1) "La capacidad del personal participante en la investigación.
2) La asequibilidad y corrección de las fuentes de información necesarias.
3) El nivel de operacionalidad.
4) El tiempo que se requiere.
5) El coste: humano, social y económico.
6) Las limitaciones deontológicas."

Pero la práctica real de la investigación suele ser muy dispar. De acuerdo con Ibáñez (1989), el proceso de selección se simplifica bastante. El investigador social suele elegir, sin pensar demasiado en la elección, la *técnica* que tiene más a mano, bien por:

a) Razones *personales:* uno es experto en esa técnica.
b) Razones *organizativas:* uno trabaja en una organización constituida para trabajar con esa técnica.
c) Razones *institucionales:* uno pertenece a una institución interesada en vender esa técnica.

Sea como fuere el proceso de selección, el investigador deberá justificar (en su *proyecto* de investigación) por qué eligió unas *estrategias* y *técnicas* con preferencias a otras. La exposición de razones suele fundamentarse en tres aspectos esenciales: los *objetivos,* las *características,* y las *condiciones de realización* de la investigación.

Además, téngase presente que, en el curso de la investigación, todo *diseño* inicial puede verse alterado, ante las dificultades que pueda suscitar su puesta en práctica. De ahí la conveniencia de que el investigador prevea, en la medida de lo posible, alternativas al *diseño* original de la investigación.

3.1.4. La factibilidad de la investigación

La fase final del *proyecto* incluye la exposición de las condiciones mínimas requeridas para hacer viable, o factible, la investigación diseñada. Estas condiciones o requisitos cabe resumirlos en los apartados siguientes:

a) Fuentes

En el *proyecto de investigación* suele adjuntarse una *bibliografía* básica, especializada en el tema que constituye el eje de la investigación, tanto en su vertiente teórica como en la empírica y metodológica.

Consistirá en una selección de *obras clave* y de publicaciones actualizadas.

b) Recursos (materiales y humanos) disponibles

El equipo investigador acredita su *capacidad* (experiencia) y *medios* para cumplir el *proyecto de investigación.*

Al *proyecto* se añade, al menos, el *curriculum* del investigador principal o de los integrantes del equipo investigador.

En el *currículum* ha de incluirse el detalle de la *experiencia* tenida en el área de conocimiento donde se inserta el *problema* o *tema de investigación.*

c) Otros recursos (materiales, económicos y humanos) necesarios

Si se está solicitando financiación económica de la investigación, en el *proyecto* también habrá de especificarse la *cantidad* necesaria para cada una de las *partidas* siguientes:

1) *Personal* (disponible y el que habría de contratarse al efecto).
2) *Equipo material inventariable.*

3) *Material fungible:* fotocopias, teléfono, correos.
4) *Dietas y desplazamientos.*

d) Planificación del tiempo de ejecución de la investigación

El investigador también ha de concretar la *cronología de tareas*. Ello supone delimitar la duración de cada fase de la investigación. En función de esto, habrá que precisar los *plazos de entrega parciales* (si los hubiese), y fijar la *fecha de entrega final.*

Existen diversas *técnicas* para el *cálculo de los tiempos de duración* de cada una de las etapas del *proyecto de investigación,* su *coste,* y las posibles *variaciones* que puedan acontecer.

Una de las técnicas más populares es el *método PERT (Program Evaluation and Review Technique* o *Program Evaluation/Research Task).* Básicamente, consiste en un *grafo* en el que se especifica cada una de las *tareas*, cómo dependen unas de otras, en qué *tiempo* se realizarán y qué *probabilidad de retraso* existe en su ejecución. Para su cálculo normalmente se pregunta a los responsables de cada una de las tareas el tiempo de duración habitual de su trabajo.

Todos estos detalles que conforman la última fase del *proyecto de investigación* son esenciales, en cuanto que ayudan al buen desarrollo de la investigación proyectada. Por lo que no hay que minosvalorar el esfuerzo que se dedique a su concreción. Como bien apunta Hakim (1994: 157), "quizás el error práctico más común consiste en subestimar el presupuesto, en tiempo y dinero, requerido para un proyecto".

En la Figura 3.1 se resume cada una que las *fases* que forman un *proyecto de investigación.* Como puede verse, cada *fase* viene determinada por la etapa que la precede. Asimismo, los *objetivos* de la investigación, junto con los *recursos* y el *tiempo disponible* para su realización, se hallan presentes en cada una de las fases del *proyecto de investigación,* marcando su configuración final.

Por último, hay que insistir en la idea de que todo *proyecto de investigación* representa, como su nombre indica, un "proyecto". Su puesta en práctica puede, por tanto, llevar consigo algunas modificaciones en una o en varias de las partes que lo conforman. De la pericia e ingenio del investigador (además de los medios a su disposición) dependerá, en buena medida, que la investigación finalice con éxito.

3.2. Diseños y estrategias de investigación

En el campo específico de los *métodos* y las *técnicas* de investigación social, a menudo se habla, indistintamente, de *métodos* o *técnicas.* Sin embargo, son cada vez más los autores que reclaman la distinción entre estos dos términos; algunos desde hace tiempo, como Greenwood (1973) o Bulmer (1984), otros en fechas más recientes, como Bryman (1995).

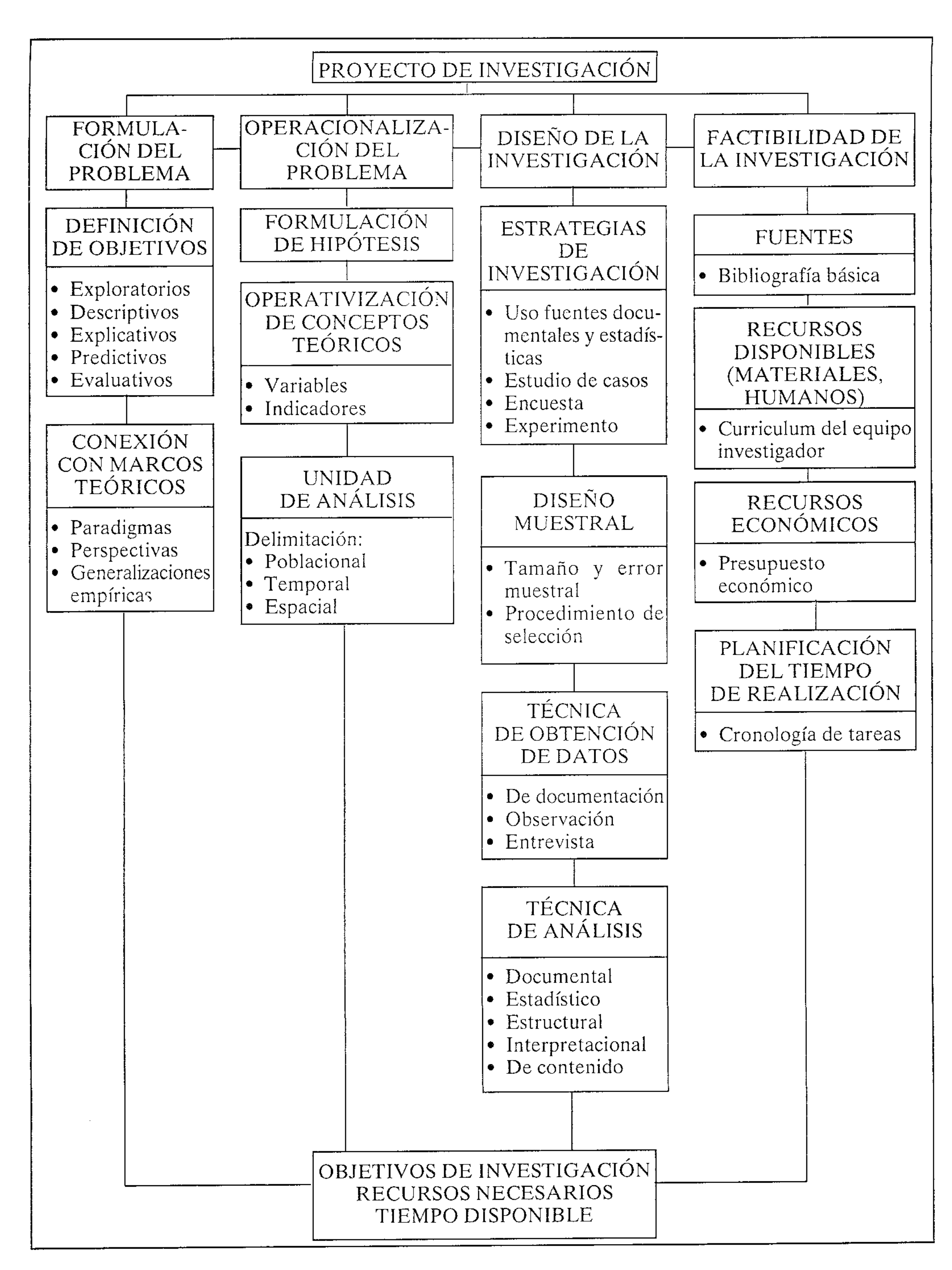


Figura 3.1. El proyecto de investigación.

Greenwood (1973), por ejemplo, emplea una analogía que coincide con el uso que en este manual se da al vocablo *estrategia:* "el método es a la técnica lo que la estrategia a la táctica. Es decir, que la técnica está subordinada al método, es un auxiliar de éste" (Greenwood, 1973: 107).

Bulmer (1984, 1992: 4-5) va más allá y propone diferenciar entre:

a) *Metodología general:* "el estudio lógico y sistemático de los principios generales que guían la investigación sociológica".

b) *Estrategia (o procedimiento) de investigación:* "manera en que un estudio empírico particular es diseñado y ejecutado".

Incluye el *enfoque* que se dará a la investigación, el *tipo de diseño* que se utilizará, y la combinación particular de las técnicas de investigación que se empleará.

c) *Técnicas de investigación:* "operaciones manipulativas específicas de búsqueda de datos".

Partiendo de esta clasificación, conviene, no obstante, barajar un concepto que ocupe un orden semántico superior al de *estrategia*. Se trata del concepto de *diseño*. En él se concentran –como ya se ha expuesto– las tareas directivas (o de ingeniería) de una investigación. Concretamente, la selección de la *estrategia* o *estrategias* a seguir y, en función de ellas, la elección de las *técnicas* (de *recogida de información* y de *análisis)* a utilizar.

Todo ello estará determinado por los *objetivos del estudio* y el *marco teórico de referencia.*

3.2.1. Pluralidad de estrategias de investigación

Como es constante en el ámbito de conocimiento de la *metodología*, no existe unanimidad entre los autores a la hora de diferenciar las diversas *estrategias de investigación.*

En la literatura consultada se observa la inexistencia de coincidencia plena, respecto a cuáles son las principales *estrategias de investigación social.* Si bien, no puede afirmarse que la discrepancia entre los autores sea abismal. Para ilustrar este extremo, el Cuadro 3.2 resume algunas *tipologías de estrategias de investigación*, seleccionadas y ordenadas por fecha de edición.

Bulmer (1992: 13) reconoce que su clasificación no es exhaustiva, aunque sostiene que los cuatro tipos diferenciados (a los que se refiere indistintamente como *"estilos"* y *"estrategias"* de investigación –al igual que hiciera posteriormente Yin (1989)– "cubre, probablemente las principales aproximaciones empleadas en nueve de cada diez monografías sobre investigación sociológica".

Bulmer (1984) destaca, como *estrategia de investigación*, la que él denomina *"métodos y fuentes del historiador";* en cambio, desestima otras estrategias importantes, como el *"experimento",* al contrario de Yin (1989).

CUADRO 3.2. Principales estrategias de investigación social, según autores.

BULMER (1984)	YIN (1989)	MARSHALL Y ROSSMAN (1989)	BREWER Y HUNTER (1989)
Encuesta. Medidas no reactivas (datos preexistentes). Métodos y fuentes del historiador. Procedimientos interpretativos (trabajo de campo).	Experimento. Encuesta. Análisis de datos de archivos. Historia. Estudio de casos.	Experimento y cuasiexperimento. Encuesta. Análisis de datos de archivos. Historia. Estudio de casos. Estudio de campo. Etnografía	Trabajo de campo. Encuesta. Experimentación. Investigación no reactiva. Aproximación multimétodo.

En su monografía sobre la investigación mediante *estudio de casos,* Yin (1989) diferencia cinco *estrategias de investigación social,* sumando el *"experimento"* a la clasificación anterior. Su formación en historia y en psicología experimental le lleva a enfatizar la *experimentación* y los *métodos del historiador.* No obstante, la *estrategia* que este autor destaca más es la denominada *"estudio de casos".*

A las *estrategias* apuntadas por Yin (1989, e. o. 1984), Marshall y Rossman (1989) añaden el *"estudio de campo"* y la *"etnografía".* Si bien, estos autores consideran a las *etnografías* como casos especiales de *estudios de campo.*

Por último, Brewer y Hunter (1989) resumen a cuatro las *estrategias de investigación social* posibles: *"trabajo de campo", "encuesta", "experimentación"* e *"investigación no reactiva"* (haciéndose eco de la publicación de Webb *et al.,* de 1966, *Unobtrusive Measures: Nonreactive Research in the Social Sciences).* A ellas suman una quinta *estrategia:* la *"aproximación multimétodo".* Ésta constituye la combinación de dos o más *estrategias de investigación.*

Los demás autores también partieron de una concepción de las *estrategias de investigación* como complementarias, abogando por una visión pluralista de las mismas. Sin embargo, no destacaron la *triangulación* o *"aproximación multimétodo"* como una *estrategia* específica y alternativa de investigación.

A partir de estas clasificaciones, propongo la siguiente *tipología de estrategias de investigación social:*

a) *Uso de fuentes documentales y estadísticas.*
b) *Estudio de casos* (etnográfico, biográfico, otros cualitativos y los cuantitativos).
c) *Encuesta.*
d) *Experimentación.*
e) La *triangulación* o *aproximación multimétodo.*

La *estrategia* que Bulmer (1984) llama *"métodos y fuentes del historiador",* y que tanto Yin (1989), como Marshall y Rossman (1989), resumen en la denominación de *"historia"*, optó por subsumirla en una *estrategia de investigación* más netamente sociológica que denomino: *uso de fuentes documentales y estadísticas.*

Ésta constituye una *estrategia básica* en cualquier *diseño de investigación* e incluye a las otras designaciones: *fuentes de datos preexistentes, análisis de datos de archivos, investigación no reactiva.*

Es una *estrategia básica* porque –como ya se expuso en la sección primera de este capítulo–, antes de proceder al diseño de cualquier investigación, resulta necesario efectuar una *revisión bibliográfica* exhaustiva sobre el tema concreto de estudio y el área de conocimiento específica donde éste se ubique. Ello favorece la familiarización con el *tema de investigación,* además de contribuir a la consecución de un buen *diseño de investigación.*

Las *estrategias* diferenciadas dentro de la *metodología cualitativa ("trabajo de campo", "etnografía", "estudio de casos")* resuelvo enmarcarlas en la denominación tradicional de *estudio de casos.*

El *estudio de casos* constituye una *estrategia de investigación* de orientación diferente a la *encuesta,* la *experimentación,* y al *uso de documentos y estadísticas;* si bien, puede hacer uso de las mismas *técnicas de recogida y de análisis* de la información.

Tradicionalmente, el *estudio de casos* se ha ubicado en la *metodología cualitativa.* Pero, como apunta Yin (1989: 24-25), los *estudios de casos* "pueden basarse enteramente en evidencia cuantitativa y no necesitan incluir siempre observaciones directas y detalladas como fuentes de evidencia".

Años más tarde, Bryman (1995: 170) insiste también en que "no todos los estudios de casos pueden ser adecuadamente descritos como ejemplos de investigación cualitativa, ya que algunas veces realizan un uso sustancial de métodos de investigación cuantitativa".

Una persona, una familia, una comunidad, acontecimiento o actividad, puede constituir el *caso*, que será analizado mediante una variedad de *técnicas de recogida de información.*

El *estudio de casos* puede ser *único* o *múltiple,* depende del número de *casos* que se analicen (desde uno hasta cientos). Por otra parte, el análisis que se haga de los *casos* puede también variar. Ello determina el carácter que adquirirá el *estudio de casos etnográfico, etnometodológico, biográfico* e, incluso, *cuantitativo.*

Cada una de las *estrategias* de investigación social referidas posee unas ventajas y unos inconvenientes. La elección entre ellas estará sobre todo determinada por:

a) El *objetivo* del estudio.
b) El *control* que el investigador desee ejercer en el desarrollo de la investigación.
c) La *orientación* de la indagación, por fenómenos actuales (en un momento concreto), o hechos del pasado (históricos).

A estos aspectos se sumarían los relacionados con la *factibilidad de la investigación* y las *preferencias del investigador* (ya expuestos en la sección 3.1).

Sin embargo, como el clásico *dictum* de Trow (1957) recomienda, y del que se hacen eco numerosos autores –entre los cuales se encuentra Bulmer (1992: 15)–, el *problema de estudio* es lo que dicta "los métodos de investigación", y no sólo las preferencias por un *estilo o estrategia de investigación.*

Así, por ejemplo, al *estudio de casos* suele describírsele como una *estrategia de investigación* que:

a) Se adecua a cuestiones del "cómo" y el "por qué".
b) En situaciones donde el investigador ejerce poco control sobre los acontecimientos.
c) El foco de atención se encuentra en un fenómeno contemporáneo, dentro de algún contexto de la vida real. Aunque, también, puede obtenerse información de hechos del pasado en los *estudios de casos biográficos.*

Pero la aplicación del *estudio de casos* no sólo se limita a estas situaciones. De acuerdo con Hakim (1994: 61):

> "Los estudios de casos son, probablemente, los más flexibles de todos los diseños de investigación. En el nivel más sencillo, proporcionan explicaciones descriptivas de uno o más casos. Cuando se utilizan de una forma intelectualmente rigurosa para lograr aislamiento experimental de factores sociales seleccionados, ofrecen la fuerza de la investigación experimental en entornos naturales."

Los *estudios de casos* más famosos han sido tanto *descriptivos* (como es la famosa investigación de Whyte, *Street corner society,* en la que se hace uso de la *técnica de la observación participante)* como *explicativos* (sirva de ejemplo la obra publicada por Allison en 1971, *Essence of decision making. Explaining in the Cuban Missile Crisis).*

Pero, el *estudio de casos* también se adecua a los otros *objetivos de investigación: exploratorios, predictivos* y *evaluativos.* De él, se ha llegado incluso a afirmar que ocupa un lugar distintivo en la *investigación evaluativa* (Yin, 1989; Stake, 1994).

Las otras *estrategias de investigación* se detallarán en los capítulos siguientes: el *uso de fuentes documentales y estadísticas,* en el Capítulo 6; la *encuesta,* en el Capítulo 7; y la *experimentación,* en el Capítulo 8. Para una profundización en el *estudio de casos* remito al manual anteriormente citado de Miguel S. Valles, en esta misma colección.

A modo de conclusión de este subapartado, repárese en la conveniencia, a veces incluso exigencia, de una concepción pluralista de las diversas *estrategias de investigación social.*

Una misma *estrategia* puede permitir la consecución de distintos objetivos de investigación; y, a la inversa, un mismo objetivo puede suponer la aplicación de dos o más *estrategias de investigación* (como ya se expuso en el subapartado 2.1.2, dedicado a la *triangulación* o articulación de *estrategias* en una misma investigación).

Asimismo, para cualquier *objetivo de investigación* existe un amplio abanico de *estrategias* y técnicas a las que recurrir: "No importa lo que quieras averiguar, probablemente existan muchas formas de hacerlo" (Babbie, 1992: 89)."

3.2.2. *Tipologías de diseños de investigación*

Como en las *estrategias,* también existen diferentes clasificaciones de *diseños de investigación.* Depende de cuáles sean los *objetivos del estudio,* el *marco temporal* y el *marco contextual* de la observación (entorno natural o entorno artificial), primordialmente.

Tal vez una de las tipologías más conocidas e influyentes en la investigación social sea la que Campbell y Stanley (1970) publicaran en 1966 (posteriormente revisada por Cook y Campbell en 1977 y 1979). Pero, no es ésta la única tipología diferenciada.

El abanico de *diseños* no se limita a la clasificación originaria que Campbell y Stanley hicieran desde la lógica experimental. Puede ampliarse más, en función de cómo se trate la variable *tiempo (diseños seccionales* o *transversales* y *diseños longitudinales),* y de cuál sea el *objetivo* fundamental en la investigación. De ahí que en este manual se distingan tres *tipologías de diseños.* Éstos se resumen en el Cuadro 3.3.

• *Diseños preexperimentales, cuasiexperimentales* y *experimentales*

En su famoso libro sobre el *diseño experimental (Experimental and Quasi-experimental Designs for Research),* editado por vez primera en 1963, Campbell y Stanley diferenciaron tres tipos principales de *diseños de investigación:* los *diseños preexperimentales,* los *cuasiexperimentales,* y los *experimentales.* Esta tipología de *diseños* responde a los siguientes criterios de clasificación:

a) Cómo se seleccionan las *unidades de observación.*
b) El *número de observaciones* realizadas.
c) El grado de *intervención del investigador* en el proceso de investigación.
d) El *control* de posibles variables explicativas alternativas a las variables analizadas *(validez interna).*
e) Posibilidad de *generalización de los resultados* de la investigación a otros contextos espaciales y temporales *(validez externa).*

CUADRO 3.3. Clasificación de los diseños de investigación.

- Según el *grado de cumplimiento de los supuestos de la experimentación:*

  1. Diseños preexperimentales o correlacionales.
  2. Diseños cuasiexperimentales.
  3. Diseños experimentales.

- Según el *tratamiento de la variable tiempo:*

  1. Diseños seccionales o transversales.
  2. Diseños longitudinales.

     — De tendencias.
     — De cohorte.
     — Panel.

- En función de *los objetivos de la investigación:*

  1. Diseños exploratorios.
  2. Diseños descriptivos.
  3. Diseños explicativos.
  4. Diseños predictivos.
  5. Diseños evaluativos.

     — Evaluación del impacto.
     — Evaluación del proceso.
     — Valoración de necesidades.
     — Evaluación mediante análisis de sistemas.
     — Análisis coste-beneficio.
     — Evaluación de conjunto.

1) *Diseños preexperimentales* o *correlacionales*

Entre sus *características* distintivas destacan:

a) La *ausencia de manipulación* de las variables intervinientes en la investigación. El investigador se limita a la observación del fenómeno que analiza, sin introducir ninguna modificación o alteración en el mismo.
b) Se efectúa una *única medición* del fenómeno, aunque se incluyan diversos aspectos del mismo.
c) *Falta de control* de posibles fuentes de invalidación de la investigación, lo que resta poder explicativo a estos diseños.

Ahora bien, la aplicación de técnicas de *análisis multivariable* permite que los *diseños preexperimentales* (o *correlacionales)* también lleguen al establecimiento de relaciones causales, por medio del *control a posteriori;* no cuando se diseña la investigación (como acontece en los *diseños experimentales),* sino después de haberse recogido la información, en la fase de análisis.

Los *diseños preexperimentales* incluyen distintas modalidades, cuya comprensión precisa de la lectura del Capítulo 8, dedicado a la *experimentación* como *estrategia de investigación.*

a) El *diseño de un único grupo con una sola medición* (posterior al *tratamiento).*
b) *Diseño pretest-postest de un único grupo.*
c) Comparación entre un *grupo experimental y otro de control, pero sin ninguna medición previa.*

Ninguno de estos tres *diseños* cumplen todos los requisitos de la *experimentación*: la existencia de al menos un *grupo de control* (no expuesto al *tratamiento* cuyos efectos quiere medirse); la *aleatorización* en la formación de los grupos; y la *medición antes y después del tratamiento experimental.* Por esta razón los autores los denominan *"diseños preexperimentales".*

El primer diseño carece de *medición* (en la *variable dependiente)* previa al *tratamiento* (o manipulación de la *variable independiente).* Ello, unido a la carencia de un *grupo de control* que ayude a la eliminación de explicaciones alternativas al *tratamiento,* impide diferenciar los *efectos* debidos al *tratamiento* de los provocados por factores ajenos al mismo.

Los otros dos *diseños* resuelven una de esta dos salvedades, pero no las dos. En el segundo diseño (el *diseño pretest-postest de un único grupo)* se introduce la medición de la *variable dependiente* antes de la aplicación del *tratamiento;* ello permite conocer la evolución de los sujetos analizados. En cambio, en el tercer diseño se incluye el *grupo de control,* pero se carece de *medición* previa al *tratamiento.* Hecho que dificulta, igualmente, el establecimiento de *relaciones causales.*

Como ejemplo de *diseño preexperimental* o *correlacional* destaca la *encuesta* usual (no la *encuesta panel),* en la que se produce una única medición de la realidad social.

2) *Diseños experimentales*

Cabe definirlos como "observación controlada" (Caplow, 1977), gracias a una serie de actuaciones del investigador, dirigidas al control de posibles fuentes de invalidación de la investigación. Entre estas actuaciones destacan:

a) La *manipulación experimental.* En los *diseños experimentales* el investigador puede manipular *a priori* (es decir, antes de la recogida de información) las va-

riables cuya influencia en la ocurrencia de determinados fenómenos trate de medir.

b) La *formación de grupos de control,* que sean totalmente equivalentes al *grupo experimental,* salvo en la variable o *variables independientes* cuyos efectos se trata de medir.

   Los individuos se asignan, de forma *aleatoria,* a cada uno de los grupos: el *experimental* y el de *control.* De esta forma se garantiza la equivalencia inicial en la constitución de ambos tipos de grupos.

Estas actuaciones favorecen el análisis de *relaciones causales,* al cumplirse las exigencias de *validez interna:* el control de posibles explicaciones alternativas a las analizadas.

No obstante, los *diseños experimentales* presentan una importante desventaja: la mayoría de ellos carecen de *validez externa,* debido, precisamente, a la *manipulación experimental* (a la alteración que el investigador introduce en la realidad que analiza). Ello imposibilita la generalización de los resultados de la investigación.

La falta de *validez externa* también suele deberse a las características de la *muestra.* Esto acontece cuando los sujetos que participan en el *experimento* no se seleccionan al azar, entre los que constituyen el *universo* o *población* del estudio, sino que se reclutan entre los voluntarios al *experimento.* Ello, sumado al usual escaso número de integrantes en la *muestra* (no superior a los 200 casos, generalmente), limita la posibilidad de generalización de los resultados de la investigación a otros contextos distintos al experimental *(validez externa).*

Bajo la rúbrica de *diseños experimentales* se incluye una variedad de *diseños* de investigación. Estos se detallan en el Capítulo 8, dedicado a la *experimentación.* La lectura de este capítulo es necesaria para la comprensión de la lógica experimental.

3) *Diseños cuasiexperimentales*

Se hallan a caballo entre los dos polos extremos de *diseños* anteriormente referidos, participando de las características definitorias de ambos.

En los *diseños cuasiexperimentales* puede haber (o no) *manipulación experimental* de la *variable independiente* o predictora, con el propósito de comprobar su efecto en la *variable dependiente* (cuyos valores –como su nombre indica– "dependerán" de los que tome la *variable independiente).* Pero, se distancian de los *diseños experimentales* en dos aspectos fundamentales:

a) Los *diseños cuasiexperimentales* rara vez acontecen en el marco de un laboratorio, sino en el contexto de la vida real.

b) La distribución de las unidades de observación (la *muestra* del estudio) al *grupo experimental* y al de *control* no se realiza de forma *aleatoria*. Por lo que no puede garantizarse la equivalencia inicial de los grupos de comparación.

Esta última limitación puede obviarse mediante la aplicación de *técnicas estadísticas bivariables* y *multivariables* de *control a posteriori*. Ello favorece la consecución de resultados similares a los alcanzados mediante los *diseños experimentales*.

De los *diseños preexperimentales* (o *correlacionales)* se distinguen por un hecho esencial: la actuación del investigador no se limita (en los *diseños cuasiexperimentales)* a la observación; sino que, por el contrario, puede estructurar la situación de forma que facilite su análisis.

En función de estas características definitorias de los *diseños cuasiexperimentales,* puede agruparse la amplia variedad de *diseños* existentes en dos categorías generales:

a) Diseños en los cuales el investigador sí *manipula* la situación experimental, pero no existe un *grupo de control.*
b) Diseños en los que el investigador no *manipula* la situación experimental, pero sí existen *grupos de control* y *experimental* equiparables.

Campbell y colaboradores distinguen tres grandes grupos de *diseños cuasiexperimentales*, que resultan de la combinación de las distintas opciones posibles:

a) Diseños que no permiten la *inferencia causal,* bien por la ausencia de algún *grupo de control* equiparable, o bien por la inexistencia de medición de las variables antes de comprobarse la influencia de un *tratamiento* o intervención.
b) Diseños con *grupo de control* no equivalente.
c) Diseños de *series temporales* (a partir de la información extraída de fuentes estadísticas y documentales, de encuestas periódicas, o mediante otra estrategia de investigación). Lo que singulariza a esta variedad de *diseño* es la medición de la *variable dependiente* en distintos períodos de tiempo, con la finalidad de analizar su evolución.

Como esta clasificación de *diseños* se hace desde la vertiente experimental, su comprensión exige el conocimiento de los rasgos distintivos de la *experimentación* como *estrategia de investigación.* Razón por la cual se reitera en la necesidad de completar esta tipología de *diseños* con la lectura del Capítulo 8.

• *Diseños seccionales (o transversales) y diseños longitudinales*

Una clasificación distinta de los *diseños de investigación* responde a cómo se haya planificado la recogida de información: en un único momento *(diseños seccionales o*

transversales) o, por el contrario, en dos o más veces *(diseños longitudinales)*. En este último caso, se pretende analizar la evolución de los fenómenos que se investigan a lo largo del "tiempo". De esta forma, el "tiempo" pasa a ser variable esencial en este tipo de *diseños*.

1) *Diseños seccionales o transversales*

Se caracterizan por circunscribir la recogida de información a un único momento en el tiempo. El objetivo de la investigación puede ser *descriptivo* (describir las características de una población en una fecha concreta, como acontece en el Censo de Población, por ejemplo), *explicativo* (analizar, mediante una *encuesta,* las variables que inciden en la ocurrencia de un hecho o acción determinada, por ejemplo), o de otro orden. Lo que distingue a esta variedad de *diseño* es que la recogida de información se lleva a cabo de una sola vez, aunque se incluyan circunstancias temporales o contextos ambientales diferentes.

EJEMPLO DE DISEÑO SECCIONAL

Si un investigador quisiera comprobar cómo incide la situación económica de un país en los hábitos de consumo de sus habitantes, el *diseño de la investigación* sería *seccional,* si la información se recogiese en un único momento: mediante una *encuesta,* por *ejemplo,* aplicada una sola vez, en varios países seleccionados conforme a su nivel de desarrollo, y de forma simultánea.

El *diseño* sería, por el contrario, *longitudinal,* si se hiciera la misma *encuesta* en fechas diferentes, con la periodicidad que el investigador determinase en el *diseño* de la investigación. En este caso, el objetivo principal del estudio sería analizar la *evolución* de los hábitos de consumo, en tiempos y países distintos.

2) *Diseños longitudinales*

Esta modalidad de *diseño* se caracteriza por plantear el análisis del *problema de estudio* a lo largo del tiempo, con el propósito de observar su dinámica. La recogida de información se planifica, entonces, para su realización en varias fechas, que el investigador especifica en el *proyecto de investigación.*

La amplitud del período de observación y su cronología se halla relacionada con el objetivo de la investigación.

EJEMPLO DE DISEÑO LONGITUDINAL

Si el objetivo principal de la investigación fuese describir cómo la experiencia de desempleo influye en la dinámica familiar de las personas en paro, la concreción del tema y su limitación en el tiempo marcará la duración del estudio: por lo general, inferior a tres años. En todo caso, el período de investigación guardará relación con la duración media estimada de la situación de desempleo en la sociedad y momento histórico que se analice.

Por el contrario, describir la experiencia de la vejez, cómo se vive en distintas sociedades y a diferentes edades, implicaría un lapso de tiempo bastante superior. El estudio podría comenzar en el momento en que una *cohorte* de personas (o generación) cumple los 65 años de edad, y finalizar cuando sus supervivientes (de la *cohorte* escogida) alcancen los 90 años. Obviamente, conforme se avance en el tiempo, disminuirá el número de personas en la *muestra del estudio;* sobre todo, cuanto mayor sea el período de tiempo considerado. Asimismo, se podría optar por seguir:

a) Un *diseño longitudinal de cohorte,* si se parte de la cohorte general de personas que en la fecha de inicio de la investigación cumple los 65 años de edad. De ellas se extraería, preferiblemente mediante un procedimiento aleatorio, una *muestra*. Para posteriores observaciones (por ejemplo, cada cinco años), se seleccionarían nuevas *muestras* del total de supervivientes de la *cohorte* inicial.

b) Un *diseño longitudinal de panel*. Este tipo de *diseño* difiere del anterior en que la investigación se fundamenta en una única *muestra*. Son las mismas personas que conformaron la *muestra* inicial del estudio las que se analizan durante toda la investigación. El inicio sería el mismo: se extrae una muestra de la *cohorte* de personas que cumple 65 años al comienzo del estudio. Pero, para posteriores observaciones, no se elegiría una nueva *muestra,* sino que volvería a recogerse información de las mismas personas que constituyeron la *muestra* inicial de la investigación.

Habría, por tanto, varias modalidades de *diseños longitudinales*. Éstos tradicionalmente se han agrupado en los tres tipos siguientes: diseños longitudinales de tendencias, de cohorte y de panel.

• *Diseño longitudinal de tendencias*

El estudio se enfoca a la descripción de la población total, no de una parte de ella. Se trata de analizar su evolución, prestando atención a las *tendencias* y cambios de *tendencias* (de ahí su nombre) en las características investigadas.

EJEMPLO DE DISEÑOS LONGITUDINALES DE TENDENCIAS

Éste es el caso cuando se comparan los datos registrados en diferentes censos de población (como los de 1960, 1970, 1981 y 1991) para analizar su evolución; o cuando se contrastan diferentes sondeos de opinión, efectuados durante el desarrollo de una campaña electoral, con la finalidad de analizar la evolución del voto; o cuando se cotejan las respuestas emitidas a unas mismas preguntas en distintas encuestas, realizadas en fechas diferentes.

En cualquiera de estos casos, se confrontan las respuestas dadas a unas mismas cuestiones. El instrumento de medición no varía; pero sí la *muestra* del estudio. En cada observación puede analizarse una *muestra* diferente (aunque extraída de la misma población que constituye el foco de la investigación). Otras veces no se extrae una *muestra,* sino que se analiza a toda la población.

La existencia de bancos de datos, donde se almacenan distintas encuestas (atendiendo a su temática), facilita el análisis de *tendencias*. Además, en muchos centros de opinión suelen realizarse sondeos periódicos, donde se formulan las mismas preguntas que en sondeos precedentes, pero a distintas personas (aunque de características semejantes a las *muestras* anteriores). Con esta práctica se pretende determinar la influencia del *tiempo* en el cambio de *tendencias* en las respuestas a distintas *encuestas*.

- *Diseño longitudinal de cohorte*

Esta segunda modalidad de *diseño longitudinal* difiere del anterior en que el interés no se halla en la población total, sino en una subpoblación o *cohorte*.

La *cohorte* estará constituida por individuos que comparten una misma característica: la edad, habitualmente. La *cohorte de edad* puede definirse de forma amplia: personas nacidas en los años sesenta en España, por ejemplo. Pero, también de una manera más específica: personas que nacieron en 1961 o, aún más concretamente, en el mes de julio de 1961.

Aunque la *cohorte de edad* constituya la *cohorte* más típica, no es la única posible. Una *cohorte* puede constituirse a partir de un acontecimiento determinado: personas que se doctoraron en las universidades españolas en 1985; parejas que se casaron en Sevilla, en 1992, durante la Expo. Igualmente, la *cohorte* puede fijarse a partir de una experiencia concreta: mujeres que estuvieron en prisión 10 años; personas mayores de 40 años que llevan dos años en situación de paro ininterrumpido, por ejemplo.

De la *cohorte* escogida se analiza su evolución. Para ello se selecciona una *muestra* distinta de entre los individuos que conforman la *cohorte* seleccionada. Es decir, se observan distintos individuos (pertenecientes a la misma *cohorte)* en momentos diferentes.

EJEMPLO DE DISEÑO LONGITUDINAL DE COHORTE

Un grupo de investigadores desea analizar cómo varían las relaciones entre padres e hijos, a medida que estos últimos van independizándose del núcleo familiar. Para ello eligen una *cohorte* compuesta por personas que nacieron en el año 1955. Elaboran un cuestionario, formado por preguntas referidas al tema en cuestión. La encuesta deciden que se pase cada cinco años, a personas que integren la *cohorte* seleccionada.

En 1975 se extrae una *muestra* de entre aquellos jóvenes de 20 años de edad; en 1980 se extrae otra *muestra* de entre las personas de 25 años; en 1985, de entre aquéllos con 30 años; y así sucesivamente, hasta la fecha de conclusión de la investigación.

Como la *muestra* se escoge de los supervivientes de la *cohorte* inicial, su volumen se verá mermado conforme se avance en el tiempo

Pero la investigación no tiene por qué circunscribirse a una única *cohorte.* Puede compararse la evolución de *cohortes* diferentes.

EJEMPLO DE DISEÑO LONGITUDINAL DE VARIAS COHORTES

En una investigación que tenga como objetivo principal determinar cómo incide la experiencia carcelaria en la reincidencia en conducta delictiva (medida por el número de detenciones policiales o de la Guardia Civil), podría compararse la evolución de personas que estuvieron en prisión durante cinco años en dos períodos de tiempo diferentes: uno, en los años sesenta (durante el franquismo); dos, a principios de los años ochenta (durante la democracia).

- *Diseño longitudinal de panel*

En esta última variedad de *diseño longitudinal,* la atención del investigador se dirige, a diferencia de los *diseños* anteriores, a analizar la evolución de unos mismos individuos, que se eligieron al inicio de la investigación. Por lo que, no se procede a nuevas selecciones muestrales en cada fase posterior de la investigación.

EJEMPLO DE DISEÑO LONGITUDINAL DE PANEL

Un ejemplo clásico de este tipo de *diseño* es la investigación que Lazarsfeld, Berelson y Gaudet llevaron a cabo en EEUU, durante las elecciones presidenciales de 1940 (la campaña electoral Wilkie-Roosevelt): *The people's c2hoice* (1944).

El estudio tenía como objetivo comprobar la influencia de la campaña electoral en las intenciones de voto de los electores del condado de Erie (Ohio). Para dicho propósito diseñaron un cuestionario, que pasaron varias veces (entre mayo y noviembre de aquel año) a la misma muestra de electores que extrajeron (siguiendo procedimientos *aleatorios*) a comienzos de la investigación (600 personas).

El *diseño panel* es de gran utilidad para indagar en las causas del "cambio". Las personas que reiteradamente se observan son las mismas; lo que varía son las circunstancias en que se hallan tras el paso del tiempo. Esto permite conocer los factores que pueden haber contribuido al cambio que en ellos se observe.

En su contra, el *diseño panel* se enfrenta a dos problemas metodológicos fundamentales:

a) El *desgaste de la muestra* con el consiguiente aumento de la *no respuesta*.

Éste representa un problema similar al denominado "mortalidad experimental". Hace referencia a la dificultad de localizar, en fechas distintas, a las personas que forman la *muestra* del estudio. Ello puede deberse a un cúmulo de razones: la persona se niega a seguir participando en la investigación; ha cambiado de domicilio y no lo ha comunicado al equipo investigador; se encuentra enferma o ha fallecimiento, entre otras posible razones.

Este problema se halla más presente en *estudios longitudinales* que abarcan un amplio período de tiempo.

El paso del tiempo actúa de forma negativa, propiciando el desinterés de los sujetos por participar en la investigación. Esto revierte en la reducción del *tamaño muestral*. Lo que introduce un sesgo importante en la investigación: las personas que abandonan el estudio pueden diferir de aquellas que permanecen en la investigación. Ello tendrá repercusión en la *validez externa* de la investigación (su significatividad y posibilidad de generalización).

b) El proceso de medición puede suscitar *sesgos en mediciones posteriores*, por efecto del aprendizaje.

Las respuestas o actitudes que una persona manifieste en una fase de la investigación puede deberse a su actuación en fases anteriores, especialmente cuanto mayor es la proximidad entre las fechas de observación. Es decir, el hábito o el ejercicio de la investigación puede llevar a la no veracidad de las respuestas o conductas manifestadas.

A estos dos problemas fundamentales del *diseño panel* se suma uno común a todo *diseño longitudinal:* el mayor coste económico de la investigación, no sólo presente en la fase de recogida de información, sino también en la de análisis (por el amplio volumen de información que se maneja). Este problema adquiere mayor relieve, conforme aumenta el tamaño de la muestra y su diversidad espacial.

El *diseño transversal o seccional* también puede hacer viable el análisis del cambio (y a un coste económico considerablemente inferior al *diseño longitudinal),* si en el instrumento de medición (cuestionario, guión de entrevista abierta, ...) se incluyen preguntas concernientes al "pasado" de los sujetos: hechos, opiniones, actitudes. Por ejemplo, a qué partido votó en las tres últimas elecciones; cuáles eran sus actividades de ocio antes de casarse; qué opina sobre el divorcio ahora y cuál era su opinión hace cinco años.

Este tipo de estudio retrospectivo, al que Hakim (1994) se refiere como "diseño cuasi-longitudinal", presenta, igualmente, un problema metodológico importante: los fallos en la memoria. La persona entrevistada puede no recordar acontecimientos de su vida pasada, sobre todo cuanto mayor es el lapso de tiempo transcurrido y menor repercusión tuvo el evento en su vida. Asimismo, puede sí recordarlos, aunque vagamente, y ser inexactas sus respuestas.

En suma, cualquier *diseño* de investigación presenta unas ventajas y, a su vez, unos inconvenientes. La elección entre un tipo de *diseño* u otro vendrá marcada por la conjunción de los tres factores señalados en el apartado 3.1. Concretamente, los *objetivos* de la investigación, los *recursos* (materiales, económicos y humanos) y el *tiempo* que se disponga para la realización del estudio.

• *Diseños exploratorios, descriptivos, explicativos, predictivos y evaluativos*

Una última tipología de *diseños* de investigación responde, precisamente, a cual sea el objetivo principal de la investigación. En 1989, dos autoras, Marshall y Rossman, publican, en su obra *Designing Qualitative Research,* una clasificación de *diseños* fundamentada en los *objetivos de la investigación;* si bien, los limitan a cuatro: *"exploratorios", "descriptivos", "explicativos"* y *"predictivos".*

A estos objetivos genéricos habría, en mi opinión, que añadir otro fundamental: el *evaluativo*. De este modo resultaría la configuración de *diseños* de investigación que se detalla a continuación.

Sin embargo, téngase presente que esta clasificación de *diseños* no ha de entenderse como si se tratase de modalidades excluyentes. Una misma investigación puede incluir *objetivos* de diverso rango, en consonancia con las distintas fases en su desarrollo. En estos casos se estaría ante *diseños de investigación* complejos.

El estudio puede comenzar siendo *exploratorio* (si el equipo investigador dispone de escasa información sobre el objeto de conocimiento) para, posteriormente, proceder a la *descripción, explicación, predicción* y/o *evaluación*. Generalmente, los *diseños* que incluyen objetivos de rango superior suelen comprender, a su vez, objetivos inferiores como fases previas en su materialización. Tal es el caso de la *investigación evaluativa,* en cuyo *diseño* confluyen varias etapas: desde la familiarización con el *programa* o la *intervención* que se evalúa (mediante un *diseño exploratorio* y *descriptivo*) hasta, por ejemplo, la valoración de los *resultados del programa* (mediante un *diseño explicativo* que anteceda al *evaluativo).*

La interpretación de los *resultados* de cualquier *programa,* intervención, o conjunto de actividades, requiere que antes se haya descrito al *programa* y su funcionamiento real. A partir de esta descripción se procede a descubrir qué se debe al *programa* y qué cabe atribuir a la intervención de otros factores. Como afirma Alvira (1991a: 7, 10-11):

> "Evaluar es más que investigar [...]. Evaluar es emitir juicios de valor, adjudicar valor o mérito de un programa/intervención, basándose en la información empírica recogida sistemática y rigurosamente."

Por último, adviértase que no se trata de una taxonomía, ni tampoco de una escala. Podría, también, haberse resaltado otros objetivos (como el comparativo, por citar alguno). Pero se desestiman por considerar que quedan subsumidos en los objetivos anteriores.

1) *Diseño exploratorio*

Un diseño de investigación exploratorio se lleva a cabo para cubrir alguno o varios de los *propósitos* siguientes:

a) Familiarización con el *problema de investigación* para deducir (a partir de la información reunida) qué aspectos requieren un análisis pormenorizado en indagaciones posteriores.
b) Verificar la factibilidad de la investigación y documentar los medios que se precisan para hacerla viable.
c) Comprobar qué *estrategia* (o estrategias) de investigación se adecúa más a su análisis.
   En cada *estrategia,* seleccionar la *técnica* (o técnicas) de obtención de datos y de análisis más pertinentes para futuras indagaciones (más formalizadas).

El *estudio exploratorio* rara vez constituye un fin en si mismo. "Rara vez proporcionan respuestas satisfactorias a las preguntas de investigación. Pueden sugerir los métodos de investigación que podrían proporcionar respuestas definitivas" (Babbie, 1992: 91). A ello contribuye la escasa representatividad de los casos que se analizan (ya por su volumen, ya por el procedimiento de selección muestral empleado).

2) *Diseño descriptivo*

Al igual que la exploración, la *descripción* constituye un paso previo en cualquier proceso de investigación. Antes de indagar en la explicación de cualquier evento hay que

proceder a su descripción mediante alguna o varias *estrategias de investigación (encuesta, uso de documentos y estadísticas o el estudio de casos)*. De ellas, el investigador obtendrá información que le servirá en la caracterización del fenómeno que analiza.

EJEMPLO DE DISEÑO DESCRIPTIVO

El *censo de población* es un ejemplo tradicional de indagación *descriptiva*. En él se describen características sociodemográficas básicas de una población determinada, en un momento concreto.

3) *Diseño explicativo*

Después de la *descripción* procede la *explicación:* buscar posibles causas o razones de los hechos, acciones, opiniones o cualquier fenómeno que se analice.

EJEMPLO DE DISEÑO EXPLICATIVO

Describir el perfil (o perfiles) de la población delincuente juvenil constituye un *estudio descriptivo*. Pero, analizar las *causas* que convergen en la delincuencia juvenil implica un *diseño explicativo*, en el que se midan las interrelaciones e influencias de distintas variables.

4) *Diseño predictivo*

La *predicción* forma un objetivo específico en sí mismo, aunque lleva consigo, a su vez, alguno o varios de los objetivos anteriormente referidos. Para poder predecir cuál será la evolución futura de un determinado fenómeno habrá, previamente, que proceder a su análisis en el momento presente, tanto en la vertiente *descriptiva* como en la *explicativa*.

5) *Diseño evaluativo*

De acuerdo con Babbie (1992: 346), "la *investigación evaluativa* –algunas veces llamada *evaluación de programas*– se refiere más a un propósito de investigación que a un método de investigación específico". Se configura como una forma de *investigación*

aplicada definida por la "aplicación sistemática de procedimientos de investigación social para asegurar la conceptualización, diseño, realización y utilidad de programas de intervención social" (Rossi y Freeman, 1991: 5).

En el *diseño evaluativo* se busca la aplicación de procedimientos de investigación *(estrategias, técnicas)*, de manera sistemática y rigurosa. Con ello se pretende alcanzar conclusiones *(válidas y fiables)* sobre la efectividad del *programa* (o conjunto de actividades específicas) cuya actuación se evalúa.

Existen varias *modalidades de evaluación*, que pueden resumirse en las siguientes:

• *Evaluación del impacto*

Incluye las denominadas *evaluación de resultados*, de *efectividad* y *sumativa*. Representa una de las variedades de evaluación más populares. Persigue –como su nombre indica– el análisis del *impacto* o *efectos de un programa* para, a partir de la información obtenida, buscar su mejora (si se precisa). Ello requiere que previamente se:

a) *Definan* los criterios de éxito: los *objetivos* o *metas* que el *programa* debería alcanzar.
b) *Diferencien* los *resultados* debidos al *programa* de los causados por otros factores.
c) *Especifiquen* las condiciones bajo las cuales el *programa* resultaría más eficaz.

• *Evaluación del proceso, formativa o del desarrollo*

Tiene como objetivo fundamental la *descripción del programa:* en qué consiste y cómo funciona, en la teoría y en la práctica. De ahí su complementariedad con la modalidad de evaluación anterior. La *evaluación del proceso* contribuye a:

a) La comprobación de si el *programa* se está realizando en conformidad con los planes originales.
b) Conocer las causas de su éxito o fracaso.

Para ello, es necesario que la *evaluación del proceso* incluya –siguiendo a Krisberg (1980)– los siguientes aspectos:

a) Las condiciones y los supuestos que definan, operativa y conceptualmente, los rasgos del *programa* (supuestos teóricos, características organizativas, históricas).

b) Los criterios y los procedimientos seguidos en la selección de los clientes o participantes en cada una de las alternativas del programa.
c) Las actividades y los servicios creados para alcanzar los objetivos del *programa*.
d) Los criterios a seguir para poder determinar la efectividad del *programa*.

• *Valoración de necesidades*

En este caso, el objetivo específico de la investigación lo constituye la identificación de las *necesidades prioritarias* que el *programa* debe atender. Este tipo de valoración suele llevarse a cabo cuando el programa a evaluar resulta amplio y complejo, con un gran número de componentes que precisan revisarse (Herman *et al.*, 1987).

• *Evaluación mediante análisis de sistemas*

Se parte de la consideración del programa como un conjunto de subsistemas (personal subalterno, dirección, gerencia) relacionados y vinculados, a su vez, a otros sistemas más amplios (familia, comunidad). Se busca el análisis de sus relaciones e interrelaciones.

• *Análisis coste-beneficio*

Se estudia la relación entre los *costes del programa* y sus *resultados,* expresados (generalmente) en términos monetarios.

Difiere del *análisis efectividad-coste* en que este último no traduce los beneficios del programa a términos monetarios, sino de *impacto real.*

• *Evaluación de conjunto*

Comprende tanto las *metas del programa* como su *repercusión,* las alternativas disponibles y/o los costes de su actuación. Por lo que conforma una variedad de *evaluación de síntesis.*

La *investigación evaluativa* puede –como bien indica Alvira (1985)– desarrollar cuatro fases principales: *evaluación de necesidades, planificación del programa, evaluación formativa* y *evaluación sumativa;* solamente una, o varias de ellas, o incluso todas conjuntamente.

Asimismo, el *diseño de una investigación evaluativa* puede llevar consigo la aplicación de una única *estrategia de investigación* (el *experimento* se muestra como una de las más idóneas) o varias a la vez. Igualmente, la *validez* de sus hallazgos puede requerir

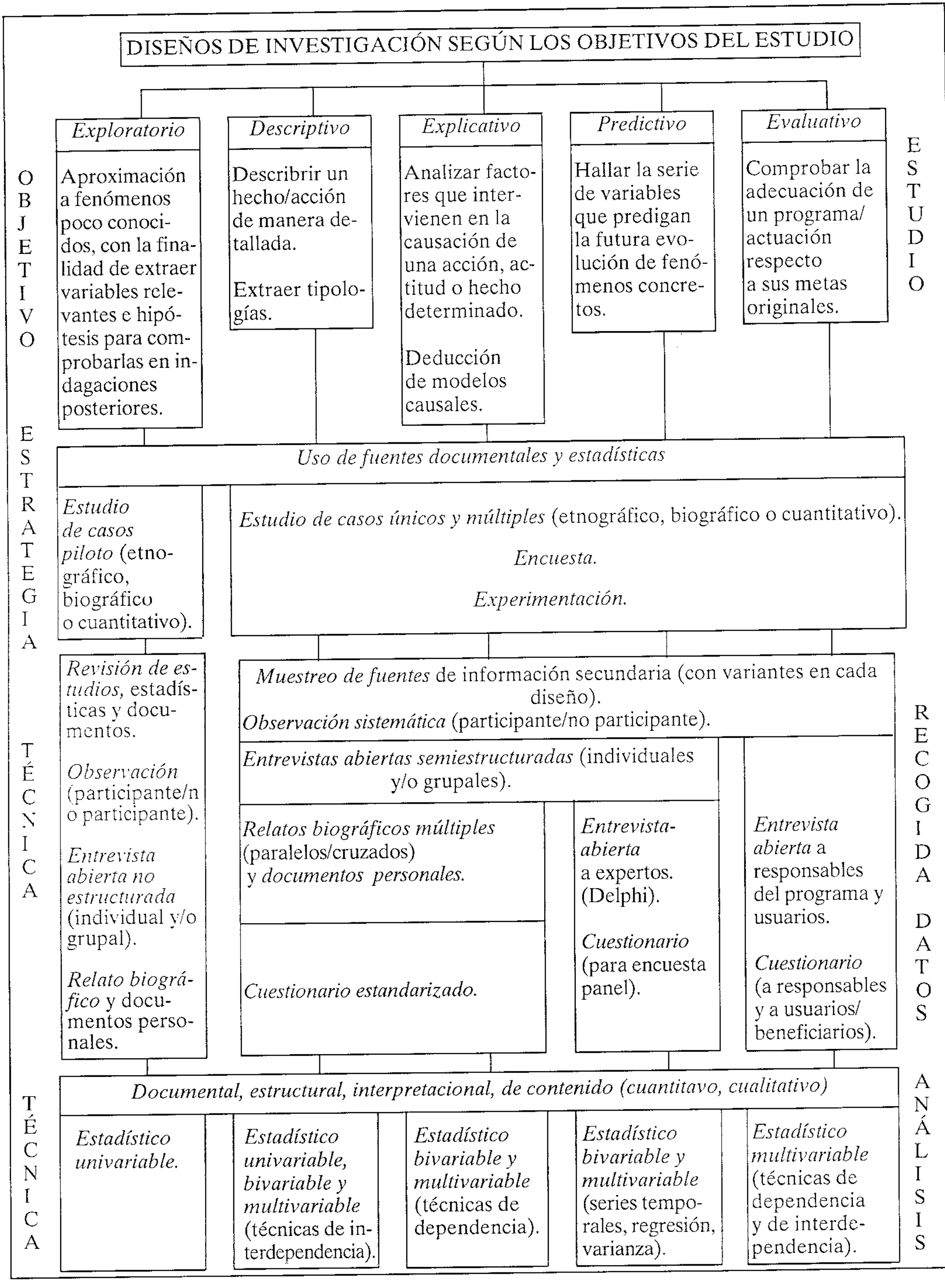


Figura 3.2. Tipología de diseños de investigación según los objetivos del estudio.

como en los otros tipos de investigación) la utilización de varias *técnicas de recogida y de análisis* de información (tanto *cuantitativas* como *cualitativas).*

La Figura 3.2 sintetiza la variedad de *diseños de investigación,* dependiendo del *objetivo principal del estudio.* Para cada tipo de *diseño* se indican distintas *estrategias* y *técnicas* (de *recogida* y de *análisis)* alternativas para cumplir los *objetivos* marcados al inicio de la indagación.

La presentación esquemática no debe mal interpretarse. Se trata de categorías genéricas de *diseños,* cuyo valor clasificatorio se consigue a costa de sacrificar el reflejo fiel y directo de la práctica investigadora.

Como ya se ha hecho referencia, cualquier investigación puede perseguir más de un *objetivo.* Puede comenzar siendo descriptiva, y terminar siendo explicativa, por ejemplo. Depende, primordialmente, del estado de conocimiento que el investigador tenga, además del enfoque que pretenda dar a la indagación.

3.2.3. Ejemplo detallado tomado de una investigación real

Como ilustración de lo expuesto, a continuación se describe el desarrollo de una investigación real. Se trata de la *investigación evaluativa* a la que se ha hecho referencia en apartados anteriores: Cea D'Ancona, M.ª A. (1992): *La Justicia de Menores en España*, Madrid, CIS, *Monografía,* n.° 127.

La investigación tenía como *objetivo principal* evaluar la actuación de los (ya extintos) Tribunales Tutelares de Menores en España. La delimitación del *problema de estudio* supuso la incardinación de la investigación en el marco de la *metodología de la evaluación de programas.* Se quería descubrir si el sistema de justicia de menores alcanzaba sus propósitos originarios (rehabilitar al transgresor, reeducándolo y adaptándolo a la vida social); o, por el contrario, producía *resultados* no deseados contribuyendo, de este modo, al desarrollo de la *"carrera delictiva"*, como afirman los teóricos del *etiquetamiento*.

A tal fin se opta, primero, por efectuar una *evaluación del proceso* (o *formativa),* que describa el sistema de justicia de menores en España (en la teoría y en la práctica), a nivel normativo y organizacional. Una vez conocido su funcionamiento real (primer objetivo), se comprueba su efectividad o *impacto* (segundo y principal objetivo de la investigación).

Ambos objetivos se ajustan a dos modalidades de *evaluación:* la *evaluación del proceso (o formativa)*, y la *evaluación del impacto (o sumativa).* A partir de ambas evaluaciones, se deducen los aspectos más necesitados de reforma en la justicia de menores (tercer objetivo específico de la investigación).

Para el cumplimiento de los objetivos marcados, se siguieron distintas fases. El primer paso fue el habitual acopio de *referencias bibliográficas,* no sólo relativas a la jurisdicción de menores y a la delincuencia juvenil, sino también de aspectos metodológicos concernientes a la *evaluación de programas* y las *técnicas de análisis de datos (bivariables* y *multivariables)* para, con posterioridad, poder abordar la parte empírica de la investigación.

A partir de la *revisión bibliográfica,* se concretan los *objetivos* y el *marco teórico* de la investigación. Para alcanzar cada uno de los *objetivos específicos* se procedió de la forma siguiente:

• *Evaluación del proceso*

Lo primero fue documentarse sobre el fundamento y el origen de la justicia de menores, instituciones y agentes colaboradores. La lectura comienza con el texto refundido de la legislación de los tribunales tutelares de menores aún vigente en España –en el momento de la investigación– (la LTTM de 1948), y prosigue con varios estudios de derecho comparado de menores en España, Alemania Federal, Francia, Inglaterra, Italia y Estados Unidos.

De estos estudios se extrajeron los puntos que diferencian y asemejan a los diversos sistemas legislativos en materia de creación, organización, competencia y funcionamiento de los tribunales e instituciones auxiliares; además de su concepción de la conducta delictiva y de la minoría de edad penal, como fundamento y razón de ser de la jurisdicción de menores.

Después se procede a analizar la práctica real del Tribunal Tutelar de Menores de Madrid (cómo opera en realidad). Los criterios seguidos para determinar la *muestra* de la investigación fueron los siguientes:

a) Como el objetivo de la investigación era evaluar la intervención del Tribunal, la población cuya actividad delictiva "registrada" resultaba episódica y, por tanto, no relevante (al menos hasta la mayoría de edad penal) no es de interés. En cambio, sí interesan aquellos jóvenes que, bien por sus características personales, bien por la entidad de los hechos cometidos (o denunciados), quedaban bajo la tutela permanente del Tribunal (ya como primer *tratamiento,* ya después de medidas leves, como la *amonestación).*

b) La *evaluación del impacto* de un *programa* requiere que la intervención del *programa* haya concluido, y no en fecha reciente. En estudios de delincuencia, el plazo de tiempo normal para medir la reincidencia es dejar transcurrir al menos dos años después de haber finalizado el *tratamiento* aplicado. Por esta razón, se excluyen del estudio los expedientes de menores abiertos o cerrados con posterioridad a 1986 (fecha de comienzo de la investigación, que duró cuatro años).

c) La *muestra* debía de ser amplia y lo más representativa posible de la población de la que se extrae, con objeto de que los resultados del estudio puedan hacerse extensivos a la población. Asimismo, debía abarcar los diversos momentos históricos vividos en fechas próximas y que, probablemente, habrían repercutido en la política de tratamiento y prevención de la delincuencia juvenil. Todo ello contribuye a la *validez externa* del *diseño de investigación.*

Con estos requisitos en mente, se consultaron las estadísticas propias del Tribunal de Menores de Madrid, referentes a su actuación desde 1975 hasta 1983 (de fechas posteriores se carecía de información estadística). En las estadísticas se observaron cambios de tendencia en la política de tratamiento de la delincuencia desde 1975. Por lo que se incluyeron todos los casos con expedientes abiertos en 1975, 1977, 1979, 1981 y 1983. Éstos debían también reunir la condición de haber estado bajo tutela del Tribunal con anterioridad a 1986. La *muestra* final de la investigación la integraron 715 expedientes de menores que habían estado en libertad vigilada y/o internamiento antes de 1986.

De los distintos informes adjuntados al expediente (de la Policía, del Servicio de Orientación Psicopedagógico del Tribunal, del delegado de Libertad Vigilada y del Centro de Reforma o de educación, además de las diligencias previas y el *acuerdo* del Tribunal) se obtuvo información para describir:

a) El *cliente* del *programa:* las características sociodemográficas y delictivas del menor de reforma. La aplicación de *técnicas de análisis multivariable* (de *conglomerados, discriminante* y *factorial confirmatorio)* contribuyó a la diferenciación de tres tipos de menores, rompiéndose el estereotipo del menor delincuente común.

b) El *proceso* de aplicación de la norma legal, desde la instrucción del procedimiento penal hasta el *acuerdo* de medida. Se contrasta la información obtenida del expediente (lo que "es"), con lo dicho al respecto en la normativa legal vigente (lo que "debería ser").

En concreto, se analizaron los distintos criterios seguidos por el Tribunal en la adjudicación de *tratamientos* a diferentes menores (a partir de la instrucción del procedimiento y del *acuerdo* tomado finalmente por el juez).

La aplicación del análisis *logit* permitió obtener las variables más relacionadas con la probabilidad de imponer al menor los distintos *tratamientos* (medidas leves, libertad vigilada e internamiento).

c) *Aplicación del tratamiento.* La *evaluación del proceso* concluyó con la descripción de cómo se ejecuta realmente el *programa:* si los diferentes *tratamientos* acordados por el Tribunal se aplicaron en conformidad con las normas establecidas. Para ello se contrastó, nuevamente, la información extraída del expediente con las disposiciones expresas en el texto legal (LTTM de 1948).

• *Evaluación del impacto*

Toda *evaluación del impacto* debe, primero, definir claramente cuáles son los criterios de éxito y de fracaso del *programa;* y, segundo, diferenciar aquellos que se deben al programa de los causados por otros factores (McQuay, 1979).

Definir los criterios de éxito de la intervención judicial no resulta difícil. Habrá éxito siempre y cuando el *tratamiento* proporcionado consiga prevenir la *reincidencia* en

conducta delictiva, logrando la reinserción del menor en la sociedad. El problema surge en su operacionalización.

La *no-reincidencia* hay que medirla tanto durante el tiempo de aplicación del *tratamiento* como al menos dos años después de su conclusión. Del expediente del menor se extrajo el número de detenciones policiales registradas antes y durante el período de *tratamiento*. Para obtener información de detenciones posteriores al tiempo de *tratamiento*, se acudió a otras instancias: la Dirección General de la Policía y la Dirección General de Instituciones Penitenciarias.

Uno de los problemas más difíciles de resolver cuando se evalúa el *impacto* de un *programa* consiste en determinar las causas a las que se deben los *resultados* observados (no siempre atribuibles al *programa*). A veces será factible seguir un *diseño experimental*, siempre que puedan formarse grupos aleatoriamente y cuyos miembros difieran sólo en el *tratamiento* que reciban. Pero esto no siempre resulta viable. Entonces, habrá que recurrir a técnicas estadísticas de *control a posteriori* –de gran utilidad cuando no se dispone de ningún grupo de control (equivalente o no)– si quieren obviarse posibles hipótesis alternativas al *tratamiento*.

El *diseño* utilizado con mayor frecuencia en *investigaciones evaluativas* del sistema de justicia es el *diseño de un grupo de control no equivalente:* igualación de los miembros de cada uno de los grupos sólo respecto a algunas de las variables. Este *diseño*, complementado con *técnicas de análisis multivariable*, consigue –siguiendo a Crizzle y White (1980: 270)– resultados equivalentes a los que se obtendrían con el *diseño experimental*.

Para comprobar la eficacia de dos de los *tratamientos* (libertad vigilada e internamiento), primero se dividió a la población en tres grupos, de acuerdo con el *tratamiento* que hubiesen recibido: sólo libertad vigilada, sólo internamiento, y ambos *tratamientos*. Éstos no constituían grupos homogéneos (aunque pudieran asemejarse en algunos rasgos), ni habían sido formados *a priori*, de forma aleatoria. Los sujetos pasaron a uno u otro *tratamiento* en función de sus características sociodemográficas, delictivas u otras.

En esta fase de la investigación, interesaba conocer cómo incidía cada una de las circunstancias personales, sociofamiliares y legales del joven en la *reincidencia;* además de aspectos relacionados con la aplicación del *tratamiento* (tiempo de aplicación, número de diligencias policiales y de denuncias previas, durante y posteriores al *tratamiento*, adaptación del menor y motivo del cese). Sus efectos se controlaron por el tipo de *tratamiento*.

Primero, se calculó la proporción de individuos que reincidieron en conducta delictiva (para cada una de las variables consideradas). A continuación, se aplicó, de nuevo, el *análisis logit* para obtener las variables más relacionadas con la probabilidad de reincidir.

A partir de la información recabada en cada una de las fases que forman la investigación, se procedió a la *interpretación de los resultados del programa*. Los *resultados* resultaron ser bastante negativos. Dos de cada tres menores proseguían su *carrera delictiva*, después de haber abandonado la jurisdicción de menores. La información obtenida durante la *evaluación del proceso* ayudó a la búsqueda de:

a) *Causas* posibles del elevado porcentaje de reincidencia obtenido.
b) *Soluciones:* la modificación plena del *programa* (o conjunto de actuaciones de los Tribunales Tutelares de Menores en España).

3.3. La validez del diseño de investigación

A la hora de *evaluar* un *diseño de investigación* existen varios *criterios* a seguir. Tal vez el más fundamental es que el *diseño* se adecue a los *objetivos* principales de la investigación. Si éstos no se alcanzan, la investigación quedaría desaprobada, ya que el fin para el que se diseñó no logra satisfacerse. En la medida en que este *criterio de evaluación* no se cumpla, no procede pasar a considerar otros criterios.

Obviamente, en la consideración de si se han alcanzado los *objetivos* del estudio, también intervienen los otros dos aspectos presentes en todas las fases del *proyecto* de investigación: los *recursos* (materiales, económicos y humanos), y el *tiempo* disponible para formalizar la investigación. Estos dos aspectos son igualmente esenciales en cuanto que inciden en el buen término de la investigación y la consiguiente consecución de los *objetivos* propuestos.

Dificultades presupuestarias o problemas de tiempo para llevar a cabo el *diseño de investigación* original, pueden llevar a su modificación e, incluso, poner en peligro el logro de los *objetivos* fijados al inicio de la investigación.

Una vez que se cumple el criterio de los objetivos, se pasaría a analizar otros *criterios de evaluación*. De nuevo, la aportación de Campbell y Stanley (1970), Cook y Campbell (1977) y Reichardt y Cook (1979), merece destacarse.

Estos autores propusieron cuatro *criterios de validez* en la *evaluación de los diseños de investigación cuantitativa:*

a) Validez interna.
b) Validez externa.
c) Validez de constructo.
d) Validez de conclusión estadística.

3.3.1. Validez interna

Hace referencia a la posibilidad de establecer relaciones de *causalidad* entre *variables (dependientes e independientes),* al haber eliminado (o *controlado)* otras *explicaciones alternativas*. De ahí que la comprobación de este tipo de *validez* en un *diseño de investigación* sea prioritaria, sobre todo en los *diseños explicativos,* más que en los *descriptivos*.

De acuerdo con Campbell y Stanley (1970), la *validez interna* constituye el *"sine qua non"* de cualquier indagación empírica. Si este criterio de *validez* no se satisface,

los resultados de la investigación serían cuestionables. Siempre surgirían posibles *explicaciones alternativas* a las relaciones observadas.

EJEMPLO DE VALIDEZ INTERNA

En una investigación sobre el *rendimiento académico* (medido por la calificación obtenida en un examen), quiere conocerse qué variables inciden más en la consecución de una buena calificación. De las distintas variables analizadas se observa la existencia de *relación positiva* entre las variables *horas de estudio,* a la semana, dedicadas a la asignatura *(variable independiente)* y *calificación* obtenida en el examen *(variable dependiente):* la calificación resulta más elevada, conforme aumenta el número de horas de estudio.

Para que esta relación sea *válida,* habría que *controlar* el efecto de *terceras y cuartas variables* que pudiesen mediar en la relación observada. Cuanto mayor sea el número de *variables perturbadoras* que el investigador controle, mayor grado de *validez* adquirirá su investigación.

Ello lleva a *controlar* el efecto de otras variables en la relación observada entre las variables *horas de estudio* y *calificación.* Entre esas variables se encuentran las tres siguientes: el cociente de inteligencia, la asiduidad en la asistencia a clase, y el grado de nerviosismo durante la realización del examen.

Para que se pudiese concluir que sí existe relación entre las variables *horas de estudio* y *calificación,* tendría previamente que haberse comprobado que, indistintamente del cociente de inteligencia, o de la asiduidad en la asistencia a clase, conforme el alumno aumenta sus *horas de estudio,* alcanza una mejor *calificación académica.* En caso contrario, la relación no se sostendría. A esto se le llama *control de terceras variables (o de explicaciones alternativas).*

El *control de explicaciones alternativas* puede efectuarse *a priori* o *a posteriori.*

a) *A priori,* en el diseño de la investigación. Esto es posible sobre todo cuando la *estrategia de investigación* es la *experimentación.*

   El *diseño experimental* conlleva la formación de *grupos de control equivalentes al experimental,* salvo en el *tratamiento* o variable cuya influencia quiere comprobarse. La selección de los sujetos para constituir cada uno de los grupos sería totalmente aleatoria. Ello contribuye a neutralizar la influencia de *variables perturbadoras* que pudieran incidir en la relación.

b) *A posteriori,* en el proceso de análisis de la información.

   Este tipo de control resulta más habitual en la práctica de la investigación social (en el análisis de datos de encuesta, por ejemplo). Su consecución se logra mediante la aplicación de *técnicas de análisis bivariable* y *multivariable,* que permiten la formación de grupos de sujetos iguales en función de los valores de la variable que se controle.

En todos los grupos formados debe observarse las mismas variaciones, dependiendo de los valores de la *variable independiente* que se analice, para poderse afirmar que esta variable explica la ocurrencia de la *variable dependiente*.

En suma, el mayor o menor grado de *validez interna* de un *diseño de investigación* depende del *control* de *explicaciones alternativas* a las relaciones observadas; es decir, del número de *variables perturbadoras* cuya influencia se haya neutralizado o controlado en el proceso de investigación.

3.3.2. Validez externa

Representa la posibilidad de *generalización de los resultados* de una investigación, tanto a la *población* concreta de la que se ha extraído la *muestra*, como a otros *tiempos y contextos*. Pero ello exige que éstos participen de las características presentes en el contexto espacial y temporal observado.

En la práctica de la investigación social, aunque la *población* a analizar sea muy pequeña, por razones de economía (presupuestaria) y de tiempo, fundamentalmente, suele observarse sólo a una parte de dicha *población*. Esta parte constituye la *muestra* de la investigación, que deberá seleccionarse de manera que constituya una representación, a pequeña escala, de la *población* de la que se ha extraído. De ello dependerá la posibilidad de que los *resultados* de la investigación puedan generalizarse y hacerse extensibles a la *población* de referencia.

Además del número de casos observados, la *representatividad de la muestra* (y la consiguiente *validez externa* de la investigación) también estará subordinada al procedimiento seguido en la selección de los elementos de la *muestra*.

Se recomienda seguir, preferiblemente, *procedimientos de selección aleatorios o probabilísticos* (que se exponen en el Capítulo 5). Cuando ello no sea factible (por carecer de un *marco de muestreo* adecuado, o por otra razón), procurar que en la *muestra* escogida se hallen incluidas las diferentes variedades de casos que componen el *universo o población de estudio*.

3.3.3. Validez de constructo

Hace referencia al grado de adecuación conseguido en la medición de los *conceptos* centrales de la investigación.

Como ya se ha hecho mención, cualquier *concepto* permite distintas posibilidades de medición. Asimismo, cualquier operacionalización de un *concepto* es difícil que cubra todas las *dimensiones del concepto*. En consecuencia, habrá que procurar operacionalizar los *conceptos teóricos* lo más rigurosamente posible, al menos los *conceptos* que sean fundamentales en la investigación. Ello contribuirá a reducir la duda de si se

habrían alcanzado los mismos resultados con una *operacionalización* distinta del *concepto*.

Para dicho propósito, se recomienda efectuar una *operacionalización múltiple*, porque permite una mejor aproximación al significado real del *concepto*. De lo que se trata es de buscar una serie de *medidas* (dos o más) para cada *concepto*. Pero, antes de afirmar que los hallazgos del estudio son válidos o inválidos, dependiendo de si existe convergencia (o divergencia) entre los resultados alcanzados en las distintas *mediciones*, ha de comprobarse si, realmente, se ha medido el mismo *concepto*.

EJEMPLO DE VALIDEZ DE CONSTRUCTO

Si la *actividad delictiva* se mide combinando *datos policiales* con la información obtenida mediante *encuesta* (tanto de *encuestas de victimización* como de las denominadas de *delincuencia autodenunciada),* lo más probable es que los resultados alcanzados, mediante estos tres procedimientos, no coincidan. La razón se halla en que en cada procedimiento mide un *concepto* distinto de *delincuencia:*

a) Las *encuestas de autodenuncia* miden la actividad delictiva de personas que no han sido todavía etiquetadas oficialmente como "delincuentes". Generalmente, hacen referencia a infracciones triviales.
b) Los *informes oficiales* (de la policía, tribunales) se refieren a infracciones graves. Suelen utilizarse para analizar la actuación judicial y/o policial, más que para medir la comisión de actividades delictivas.
c) Las *encuestas de victimización* aluden a actos delictivos cometidos por otros.

En consecuencia, los datos no pueden coincidir porque las tres medidas empleadas en realidad no miden el mismo *concepto*. De ahí la insistencia en que cuando se diseñe una investigación, se procure asegurar la equivalencia de las distintas mediciones de un mismo *concepto*. Cuando esta exigencia no se cumpla, los resultados logrados de cada una de ellas no serán comparables.

3.3.4. Validez de conclusión estadística

Esta última variedad de *validez* se halla relacionada con el poder, adecuación y fiabilidad de la *técnica de análisis de datos* aplicada.

Como la práctica de la investigación social se fundamenta en la información extraída de una *muestra*, la adecuación del *tamaño muestral* (normalmente superior a 100 casos) con la *técnica analítica* utilizada (máxime en *técnicas de análisis multivariables*, que exigen una determinada proporción de casos por variables introducidas en el análisis) incide en la *significatividad estadística* de los *resultados de la investigación;* en la

posibilidad de generalización de los hallazgos obtenidos en la *muestra* a la *población* de la que ésta se ha extraído.

Además, antes de utilizar una *técnica analítica* concreta, ha de comprobarse si la información recabada satisface los supuestos exigidos para la práctica de la técnica de análisis elegida. En las *técnicas de análisis multivariable,* los supuestos más habituales son los de *normalidad, homocedasticidad* e *independencia* de los *términos de error*. Éstos se hallan relacionados con el *tamaño de la muestra* y el *nivel de medición de las variables*.

En el Capítulo 9 se contemplan distintos aspectos a considerar en el *análisis estadístico* para que la investigación cumpla requisitos de *validez*.

• *Síntesis de los criterios de validez*

La optimización de cada uno de estos *criterios de validez* en un mismo diseño de investigación no resulta una tarea fácil, en la práctica. Primar la *validez interna,* por ejemplo, puede conllevar la disminución de la *validez externa,* ya que ambos *criterios de validez* se contraponen. Esto es, la primacia del *control experimental* lleva consigo como contrapartida, la disminución de la capacidad de generalización de los resultados de la investigación. Mientras que en los *diseños explicativos* se prima la *validez interna,* en los *descriptivos* la *externa,* máxime cuando se trabaja con *muestras*.

No obstante, puede lograrse un *diseño de investigación* que consiga un grado aceptable en los cuatro *criterios de validez* señalados. Para dicha consecución, conviene tener presente las recomendaciones que figuran resumidas en el Cuadro 3.4.

CUADRO 3.4. Recomendaciones para aumentar la validez de un diseño de investigación.

Validez interna	*Validez externa*	*Validezde constructo*	*Validez estadística*
Creación de varios grupos de comparación equivalentes al de observación. Efectuar varias mediciones. Controlar todo suceso externo e interno a la investigación que puedan afectar a sus resultados.	Selección de las unidades de la muestra mediante procedimientos aleatorios. Formar grupos heterogéneos de unidades de observación que incluyan varios contextos temporales y espaciales.	Delimitación clara y precisa de los conceptos teóricos. Operacionalización múltiple de los conceptos. Empleo de varias técnicas de obtención de información.	Aumentar el tamaño de la muestra. Formar grupos internamente poco heterogéneos.

Lecturas complementarias

Alvira, F. (1994): "Diseños de investigación social: criterios operativos", en García Ferrando, M. *et al.* (comps.), *El análisis de la realidad social*, Madrid, Alianza Universidad, pp. 87-112.
Babbie, E. (1992): *The practice of social research*. California, Wadsworth Publishing Company.
Campbell, D. y Stanley, J. (1970): *Diseños experimentales y cuasiexperimentales en la investigación social*. Buenos Aires, Amorrortu.
Hakim, C. (1994): *Research design: strategies and choices in the design of social research*. London, Routledge.
Hernández Sampieri, R. *et al.* (1991): *Metodología de la investigación*. México, McGraw Hill.
Pons, I. (1993): *Programación de la investigación social*. Madrid, CIS, Cuaderno Metodológico n.° 8.

Ejercicios Propuestos

1. Escoja una investigación y describa su diseño. En la descripción destaque si el diseño cumple criterios de validez. En caso negativo, indique qué fuentes invalidan el estudio.
2. A partir de una *idea*, formule el problema de investigación. Concrete los objetivos y el marco teórico del estudio. En su exposición, cite las referencias bibliográficas consultadas.
3. ¿Qué formas existen para controlar explicaciones alternativas? ¿De qué diseños son más características?
4. ¿En qué consiste la operacionalización múltiple? Cite un ejemplo.
5. A partir de un problema de investigación, diseñe un estudio seccional y transversal (en sus tres modalidades).
6. Durante varias semanas se observa el absentismo laboral de un grupo de personas que trabajan de noche en un hospital. Con el fin de reducir dicho absentismo, se incrementa el salario del personal del hospital que trabaja en el turno de noche. A la semana siguiente, y durante un mes, se observa el absentismo. ¿A qué tipo de diseño corresponde este ejemplo? Especifique las "amenazas" a la validez del diseño efectuado.

4

LA OPERACIONALIZACIÓN DE CONCEPTOS

La *operacionalización de conceptos teóricos* constituye una fase intermedia en el *proceso de investigación*. Antecede al diseño de la indagación empírica siendo, a su vez, consecutivo a la *formulación del problema* de estudio.

De la *teoría* (o *marco teórico* que conforma la investigación) se extraen unos conceptos y proposiciones. Los *conceptos* se traducen a *términos operacionales*. De ellos se deducen unas *variables empíricas* o *indicadores* que posibiliten la contrastación empírica del *concepto* que se analice.

El término *"operacionalización"* –tomado de las ciencias naturales– es el que habitualmente se emplea para denotar los estadios implicados en el proceso de asignación de *mediciones* a *conceptos*. El presente capítulo trata, precisamente, de la práctica de la *medición* y de las dificultades que acaecen en su desarrollo.

4.1. Fundamentos y principios de la operacionalización

En el proceso global de *operacionalización* hay que diferenciar –siguiendo a Blalock (1982)– dos nociones fundamentales: la *conceptualización* y la *medición*.

a) La *conceptualización* hace referencia al proceso teórico mediante el que se clarifican las *ideas* o *constructos teóricos*. Esta clarificación ha de hacerse de manera que la definición del *constructo teórico* comprenda el significado que se le suele asignar.

b) La *medición* connota, en cambio, el proceso general que vincula las operaciones físicas de *medición* con las operaciones matemáticas de asignar números a objetos.

El proceso completo implicaría, en consecuencia, un triple nexo que relaciona los *conceptos teóricos* con las operaciones físicas de *medición,* y de éstas con los *símbolos matemáticos*.

> "Si queremos que nuestras teorías sean generalizables a través de una variedad de entornos, o con respecto a una variedad de fenómenos, obviamente tenemos que *conceptualizar* nuestras variables de forma que las proposiciones que contengan estas variables puedan aplicarse en tales entornos y fenómenos diversos" (Blalock, 1982: 29).

Ello es necesario si se pretende la comparabilidad de las *operaciones de medición*. Pero, ¿qué se entiende por *medición?*

Tal vez la definición más popular de *medición* sea la proporcionada en 1951 por Stevens. De acuerdo con este autor, la *medición* consiste en la "asignación de números a objetos o acontecimientos de acuerdo con determinadas reglas" (Stevens, 1951: 22). Esta definición clásica de *medición* no es, sin embargo, compartida por la generalidad de los autores.

Duncan (1984: 126) matiza que la asignación de números debe hacerse de modo que los números "se correspondan con diferentes grados de *cualidad* –o propiedad– de algún objeto o evento".

Otros autores, como Carmines y Zeller (1979), van más allá. Rechazan la definición de *medición* de Stevens, por considerar que se adecua a las ciencias físicas, pero no a las ciencias sociales. En su opinión, muchos fenómenos sociales son "típicamente demasiado abstractos para ser adecuadamente caracterizados o como objetos o como acontecimientos".

Fenómenos como la *eficacia política,* la *alienación,* o la *disonancia cognitiva,* por ejemplo, son "demasiado abstractos para ser considerados cosas que pueden verse o tocarse [la definición de un objeto] o, meramente, como un resultado o consecuencia [la definición de un acontecimiento]" (Carmines y Zeller, 1979: 10).

La *medición* debería, en cambio, comprenderse –siguiendo a Carmines y Zeller (1979)– como el proceso de vincular *conceptos abstractos* a *indicadores empíricos*. Comprende, por tanto, consideraciones teóricas y empíricas.

a) Desde el *punto de vista teórico,* el interés se halla en el *concepto*. Éste no es observable (y, por tanto, no es directamente medible), sino *latente*. Es decir, se halla representado por una respuesta que sí es *observable*.

b) Desde el *punto de vista empírico,* la respuesta observable constituye el centro del proceso, tanto si ésta adquiere la forma de respuesta a una pregunta incluida en un *cuestionario* o *entrevista,* o la forma de una conducta grabada en un *estudio observacional,* por ejemplo.

En suma, la *medición* se centra "en la relación crucial entre el indicador(es) fundamentado teóricamente –eso es, la respuesta observable– y el concepto(s) no ob-

servable, latente" (Carmines y Zeller, 1979: 10). El problema se encuentra, precisamente, en la concreción de esa correspondencia (que debe haber entre el sistema conceptual y el empírico), para que se alcance una *medición* válida y fiable del fenómeno que se analiza.

Los *conceptos* pueden definirse como "símbolos lingüísticos que categorizan a los fenómenos" (Phillips, 1985: 77). Con frecuencia derivan de *reflexiones teóricas* (a las que se accede mediante una *revisión bibliográfica);* otras veces, proceden de *reflexiones propias* sobre la variedad de experiencias de la vida social.

Sea como fuere, se caracterizan por representar *constructos abstractos* y, en consecuencia, no directamente *observables* (como la *"anomia"*, la *"cohesión social"*, o la *"amistad"*, por ejemplo).

Los *conceptos* también sintetizan distintas variedades en que pueden clasificarse los objetos de conocimiento. De ahí que pueda afirmarse que los *conceptos* cumplen una función fundamental de síntesis, de denominación común, que englobe a una serie de observaciones, proporcionándolas un sentido. El *concepto* de *"tristeza"*, por ejemplo, proporciona una denominación común a una serie de manifestaciones distintas: llanto, desánimo, bajo tono de voz, negativa a hablar, comer, problemas de sueño.

Los *conceptos* difieren, esencialmente, por su mayor o menor grado de abstracción (el concepto de *"poder"*, por ejemplo, connota un grado de abstracción superior al *concepto* de *"educación"*).

Si bien, la generalidad de los *conceptos* constituyen variables "latentes", "hipotéticas", no directamente "observables". Por lo que su concreción precisa de la traducción del *concepto teórico* a *indicadores,* a *variables empíricas (observables* o *manifiestas),* que midan las *propiedades latentes* enmarcadas en el *concepto.*

En toda *operacionalización* de *conceptos teóricos* se ha de partir de las siguientes consideraciones:

a) Entre los *indicadores* y el *concepto* a medir ha de haber una plena correspondencia. Los *indicadores* han de seleccionarse y combinarse de manera que logren representar la *propiedad latente* que el *concepto* representa (su existencia e intensidad), en condiciones de *validez* y de *fiabilidad.*
b) Los *indicadores* pueden materializarse en formas diversas (preguntas en un *cuestionario* o en una *entrevista abierta,* en el *registro* de una conducta observada, en *datos estadísticos* contabilizados en un censo, por ejemplo). Depende de cuál sea la *técnica de obtención de información* que el investigador haya seleccionado en el *diseño de la investigación.*
c) En la *operacionalización,* como en todo proceso analítico, se asumen unos márgenes de incertidumbre. La relación entre los *indicadores* y la *variable latente* (el *concepto),* que tratan de medir, siempre será supuesta, nunca plenamente "cierta"; se consideran aproximaciones en términos de "probabilidad".

Esto –como bien afirma González Blasco (1989: 236)– "evidentemente limita también el valor de la medida, pero es una limitación que hemos de aceptar si queremos medir".

En conclusión, los *indicadores* se emplean para cuantificar, e inferir, la existencia o inexistencia de una *variable latente* (un *concepto);* aunque, siempre en términos de "probabilidad". Se tratará, por tanto, de reducir el *error de medición* al mínimo posible.

4.2. La medición de variables: tipologías

> "Los procesos de conceptualización y de operacionalización pueden verse como la especificación de variables y los atributos que las componen" (Babbie, 1992: 140).

Por *variable* generalmente se entiende cualquier cualidad o característica de un objeto (o evento) que contenga, al menos, dos atributos (categorías o valores), en los que pueda clasificarse un objeto o evento determinado.

Los *atributos* son las distintas *categorías o valores* que componen la variable. En función de ellos se clasifica a los objetos (o eventos) en un grupo u otro. *Variables* como la *"edad"* (años cumplidos), la *"altura"* (centímetros), o el *"nivel de ingresos"* (en pesetas), toman *valores* (numéricos). Por el contrario, *variables* como *"sexo"* (varón, mujer), *"estado civil"* (soltero, casado, viudo, separado, divorciado), o *"satisfacción conyugal"* (bastante satisfecho, satisfecho, ni satisfecho ni insatisfecho, insatisfecho, bastante insatisfecho) adoptan *categorías*.

La *medición de una variable* consiste, precisamente, en el proceso de asignar *valores* o *categorías* a las distintas características que conforman el objeto de estudio. Para que la *medición* se realice adecuadamente se recomienda, al menos, cumplir tres requisitos básicos:

a) Exhaustividad

La medición de la variable ha de efectuarse de forma que ésta comprenda el mayor número de *atributos (categorías* o *valores)* posible. El propósito es que ninguna observación quede sin poder clasificarse. De ahí la sugerencia cuando se diseña un *cuestionario,* por ejemplo, de incluir la opción de respuesta *"otros"* (especialmente en aquellas preguntas en las que caben otras respuestas diferentes a las dadas en el cuestionario), y la categoría común *"no sabe/no contesta"* (dirigida a aquellos que decidan no emitir ninguna respuesta).

b) Exclusividad

Los distintos *atributos* que componen la *variable* deben ser mutuamente excluyentes. Por lo que deberán definirse de manera que cualquier observación sólo pueda clasificarse en términos de un único *atributo*.

EJEMPLO DE EXCLUSIVIDAD

Para ilustrar la necesidad de cumplir el requisito de *exclusividad* en la *medición de una variable,* a continuación se comparan dos mediciones alternativas de la variable "edad".

Si esta variable se midiese como en (A), las personas de 18, 25, 50 y 65 años de edad no sabrían en qué *atributo* clasificarse, al estar estos valores incluidos en más de un intervalo.

En cambio, si la *medición de la variable* se efectuase como en (B), este problema desaparecería, cumpliéndose así el requisito de *exclusividad.*

A) Edad		*B) Edad*	
18 y menos	1	Menos de 18	1
18-25	2	18-25	2
25-50	3	26-50	3
50-65	4	51-65	4
65 y más	5	Más de 65	5
NS/NC	0	NS/NC	0

c) Precisión

Realizar el mayor número de distinciones posibles. Ello contribuye a la consecución de una información más *precisa.* Tiempo habrá para agrupar las distintas *categorías* o *valores* de las *variables;* generalmente, después de haberse recabado la información (en la fase de análisis), a la vista de la frecuencia que presente cada *atributo* de la *variable.* Pero, por el contrario, nunca será factible desglosar los *atributos* después de la obtención de los datos.

EJEMPLO DE PRECISIÓN

Si la variable "nivel de instrucción" se midiese como en (B), se obtendría una información más precisa y detallada que si se midiese como en (A).

A) Nivel de instrucción		*B) Nivel de instrucción*	
Sin estudios	1	No sabe leer ni escribir	1
Primarios	2	Sólo sabe leer y/o escribir	2
Profesionales	3	Primarios de 1º grado	3
Bachillerato	4	Primarios de 2º grado	4
Medios superiores	5	Formación profesional	5
Superiores	6	Bachillerato	6
		Medios superiores (Esc. Univ.)	7
		Superiores (facultades, ETS)	8
		Otros	9

Existen distintas modalidades de variables. En el Cuadro 4.1 se resumen los criterios principales de clasificación de las variables.

CUADRO 4.1. Tipologías de variables según criterios de clasificación.

Nivel de medición	• Variables *cualitativas* o *no métricas*. 1) Nominales. 2) Ordinales. • Variables *cuantitativas* o *métricas*. 3) De intervalo. 4) De razón o proporción.
Escala de medición	1) Continuas. 2) Discretas.
Función en la investigación	1) Independientes. 2) Dependientes. 3) Perturbadoras. — De control. — Aleatorias.
Nivel de abstracción	1) Generales. 2) Intermedias. 3) Empíricas o indicadores.

• *Tipos de variables según el nivel de medición*

1) *Variables nominales*

Aquellas variables cuyos atributos sólo cumplen las condiciones esenciales de *exhaustividad* y *exclusividad*. Unicamente se hallan compuestas por distintas denominaciones, entre las que no puede establecerse ningún tipo de relación (de orden o de otra clase). Ello limita las posibilidades de análisis estadísticos en este tipo de variables.
A esta modalidad pertenecen las variables sexo, estado civil, nacionalidad, partido político, color del pelo, grupo sanguíneo, situación laboral; cualquier variable que indique una cualidad del objeto o evento que se analice, sin establecer ninguna graduación entre las distintas categorías que conforman la variable.

EJEMPLOS DE VARIABLE NOMINAL

Partido político		*Facultades universitarias*	
PSOE	1	CC. Políticas y Sociología	1
PP	2	CC. Económicas y Empresariales	2
IU	3	Psicología	3
CiU	4	Derecho	4
PNV	5	Ciencias de la Información	5
Otros	6	Medicina	6
		Farmacia	7
		Otras	8

Los números que se asignan a cada una de las *categorías* que forman la variable sirven para indicar la pertenencia a la *categoría*. Se asignan arbitrariamente y no denotan cantidades. Por lo que no se toman como *mediciones cuantitativas*.

2) *Variables ordinales*

Variables cuyos *atributos* participan de las características anteriormente referidas, a las que se suma la posibilidad de poderse "ordenar", en el sentido de "mayor que" o "menor que". No obstante, continúa sin poderse conocer la magnitud exacta que diferencia a un *atributo* de otro.

Las *variables ordinales* son, igualmente, *variables no métricas o cualitativas*. Expresan una "cualidad" del objeto o acontecimiento, no una "cantidad".

Como ejemplos pueden citarse las variables clase social, nivel de estudios, ideología política, satisfacción laboral, calificación académica, curso académico o cualquier otra que comprenda categorías "ordenables" (en un sentido u otro).

EJEMPLOS DE VARIABLE ORDINAL

Partido político		*Facultades universitarias*	
Extrema izquierda	1	Católico muy practicante	1
Izquierda	2	Católico practicante	2
Centro izquierda	3	Católico no practicante	3
Centro	4	No católico pero creyente de otra religión	4
Centro derecha	5	No creyente	5
Derecha	6		
Extrema derecha	7		

3) *Variables de intervalo*

Constituyen *variables cuantitativas* o *métricas*. Puede "cuantificarse" la distancia exacta que separa cada *valor* de la variable. Ello es posible gracias al establecimiento de alguna unidad física de medición estándar (años, pesetas, horas, minutos, centímetros, grados). Lo que posibilita que pueda afirmarse, por ejemplo, que la distancia que separa a aquellas personas de 15 y 16 años es la misma que la habida entre aquellos de 72 y 73 años. Esta capacidad permite la realización de la mayoría de las operaciones aritméticas, como se verá en el capítulo 9.

EJEMPLOS DE VARIABLE DE INTERVALO

Peso (gramos)		*Nº de habitantes*		*Ingresos (pesetas)*	
40 – 55	1	Menos de 2.000	1	Menos de 70.000	1
56 – 60	2	2.000 – 10.000	2	70.000 – 100.000	2
61 – 72	3	10.001 – 50.000	3	100.001 – 200.000	3
73 – 84	4	50.001 – 100.000	4	200.001 – 300.000	4
Más de 84	5	100.001 – 500.000	5	300.001 – 500.000	5
		Más de 500.000	6	500.001 y más	6

- Años cumplidos......................
- Tiempo que Vd. tarda en llegar a la facultad......................
- Puntuación en un test de inteligencia......................

4) *Variables de proporción o razón*

A las características del *nivel de intervalo* se suma la posibilidad de establecer un cero absoluto. Lo que permite el cálculo de "proporciones" y la realización de cualquier operación aritmética.

La mayoría de las *variables de intervalo* son, a su vez, *de razón* (ingresos, n.º de habitantes, n.º de veces que se asiste a un concierto, edad –los bebes tienen menos de 1 año–, por ejemplo); aunque no todas. Ello lleva a algunos autores, como Blalock (1978), a afirmar que la distinción entre *variables de intervalo* y *variables de razón* es puramente académica más que real. Una vez que se ha determinado la magnitud de la *unidad*, resulta difícil concebir la posibilidad de fijar cero unidades.

• Estos cuatro *niveles de medición* de las *variables (nominal, ordinal, de intervalo* y *de razón)* conforman una *escala acumulativa*. Cada *nivel* comparte las propiedades de los *niveles de medición* que le anteceden. De ahí que se recomiende:

a) Escoger el *nivel de medición* más elevado posible, con el propósito de poder así acceder a un mayor abanico de técnicas analíticas (a aplicar en los datos que finalmente se obtengan).
b) Tener siempre presente los objetivos de la investigación. Estos marcarán el rango de variación en la *medición:* si se requiere una información detallada o, por el contrario, genérica.

En suma, cuando el investigador tenga que medir *variables* que pueden ser a su vez *cuantitativas* o *cualitativas,* tendrá que decidir, por ejemplo, si proceder a una *medición de intervalo* u *ordinal.* Dependerá de cómo haya diseñado la investigación. Concretamente, de qué técnicas empleará para la recogida y el análisis de la información, en conformidad con los objetivos del estudio.

En general, se aconseja optar por la precisión y el detalle antes que por la generalidad. Tiempo habrá para resumir la información. Los *atributos* de la *variable* siempre podrán agruparse en *categorías* genéricas (durante la fase de análisis). En cambio, nunca podrá procederse a la inversa. Una vez recogida la información, el investigador no puede desmembrar *categorías* genéricas en *atributos* específicos.

EJEMPLOS DE VARIABLES EN DISTINTO NIVEL DE MEDICIÓN

Variables como "calificación académica" o "edad" pueden medirse comprendiendo *categorías* o *valores.* El investigador tendrá que elegir entre uno u otro según los objetivos de la investigación.

Calificación académica		*Edad*	
Nivel ordinal	*De razón*	*Nivel ordinal*	*De razón*
No presentado		Niño	0 – 13
Suspenso	0 – 4	Adolescente	14 – 18
Aprobado	5 – 6	Joven	19 – 30
Notable	7 – 8	Adulto	31 – 50
Sobresaliente	9 – 10	Anciano	51 – 65
Matrícula Honor			Más de 65

Asimismo, tendrá que decidir si es suficiente conocer, por ejemplo, si la persona es de ideología de izquierdas o de derechas *(nivel ordinal)* o, por el contrario, necesita saber a qué partido votó en las últimas elecciones *(nivel nominal).* En caso de duda, se aconseja –como ya se ha indicado– anteponer el detalle a la generalidad.

- *Tipos de variable según la escala de medición*

Un segundo criterio de clasificación de las variables responde a si en la *medición de la variable* se ha aplicado una *escala continua* o una *discreta*. Así se diferencia entre *variables continuas* y *discretas*.

1) *Variables continuas*

Aquellas variables en las que pueden hallarse *valores* intermedios entre dos *valores* dados, al conformar una escala ininterrumpida de *valores*.

EJEMPLOS DE VARIABLE CONTINUA

La variable "edad" es *continua* porque entre un año y otro caben valores intermedios. Así, entre los 12 y los 13 años hay infinitos valores: como 12 años, 8 meses y 15 días, por ejemplo. Lo mismo cabe decir de la variable "peso" o cualquier otra variable que sea *cuantitativa (métrica)*. Si bien, hay que matizar que no todas las variables de *intervalo* o de *razón* son, a su vez, *continuas*. Pueden ser *discretas*, como después se verá.



2) *Variables discretas*

Cuando en la escala de medición de la variable no cabe la posibilidad de hallar valores intermedios, comprendidos entre dos *atributos* de la *variable*.

La generalidad de las *variables* denominadas *cualitativas (nominales y ordinales)* son *discretas*. También lo son algunas *cuantitativas,* como el n.º de miembros de una familia (no puede haber 2 hijos y medio), o el n.º de coches vendidos, por ejemplo.

EJEMPLOS DE VARIABLE DISCRETA Y CONTINUA

1 2 3 4 5 6

Número de miembros de una familia *(variable discreta)*

Variables discretas	*Variables continuas*
Nº libros comprados Nº barómetros Nº mesas en un aula Entradas de cine pagadas	Nº libros leídos Temperatura atmosférica registrada Longitud de las mesas Películas vistas

• *Tipos de variables según su función en la investigación*

Las variables también difieren según el papel que cumplen en una investigación. Atendiendo a este tercer criterio de clasificación, se distingue entre *variables independientes, dependientes* y *perturbadoras (de control y aleatorias).*

1) *Variables independientes, explicativas* o *predictoras (X)*

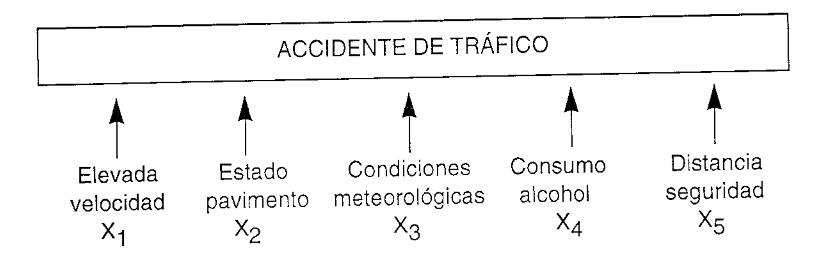
Aquellas variables cuyos *atributos* se supone que influyen en los que adopta una segunda variable (la *dependiente).* Figuran en las hipótesis de la investigación e indican posible "causas" de la variación de la variable que centra el interés de la indagación (la *dependiente* o *variable efecto).*

2) *Variables dependientes* o *criterio (Y)*

Variables cuyos atributos "dependen" –como su nombre indica– de los que adopten las *variables independientes*. Ambos tipos de variables corresponden a los objetivos de la investigación.

EJEMPLO DE VARIABLES INDEPENDIENTES Y DEPENDIENTES

Un equipo de investigadores desea determinar las variables que inciden en la ocurrencia de accidentes de tráfico. En esta investigación, el sufrir un "accidente de tráfico" actuará como la *variable dependiente* (la que constituye el objeto de análisis); mientras que las posibles "causas" de la siniestralidad serán las *independientes* (conducir a elevada velocidad, el estado del pavimento, las condiciones meteorológicas, el consumo de alcohol, la experiencia en la conducción, la edad del conductor, el guardar la distancia de seguridad).



3) *Variables perturbadoras*

En la relación entre una *variable independiente* y la *dependiente* siempre cabe la posibilidad de que existan otras variables mediando en la relación. Ello contribuye a la existencia de explicaciones alternativas que hagan espúrea la relación observada entre la *variable dependiente* y la *independiente.*

Si el efecto de esas terceras o cuartas variables se controla (bien antes *–a priori–*, o bien después de haberse recogido los datos *–a posteriori*), dejan de ser *perturbadoras* y se convierten en *variables de control.* Como ya se expuso en el Capítulo 3, en toda investigación hay que procurar controlar el mayor número de *variables perturbadoras* posible. Ello favorece la *validez interna* de la investigación.

EJEMPLO DE VARIABLES DE CONTROL

Retomando el ejemplo anterior, escójase una de las *variables independientes* barajadas: el "consumo de alcohol", por ejemplo. De la información reunida en la investigación se concluye que una alta proporción de las personas que sufrieron un accidente de tráfico presentaban un elevado índice de alcoholemia en la sangre.

Para que el investigador pueda afirmar que el consumo de alcohol (X_4) incide en la sinistralidad (Y), previamente deberá haber "controlado" el efecto de otras variables que pueden estar mediando en dicha relación. Por lo que tendrá que indagar en variables que incidan diferencialmente en personas con similares índices de alcoholemia.

Algunas de las variables que actuaban como *independientes* pueden convertirse, a su vez, en *variables de control* (cuando se analiza la influencia de otra *variable independiente);* y, a la inversa: una *variable de control* que haya mostrado incidencia pasa a ser (en la misma investigación) *variable independiente,* cuya influencia específica en la *variable dependiente* deberá cuantificarse.

(X4) Consumo de alcohol ——→ Accidente de tráfico (Y)
↑
Velocidad del vehículo.
Experiencia en la conducción.
Edad del conductor.
Condiciones atmosféricas...

Las *variables aleatorias* o *estocásticas* son, también, *variables perturbadoras,* pero de menor incidencia en el conjunto de la investigación. Su efecto en la relación entre la *variable independiente* y la *dependiente* puede reducirse siguiendo un procedimiento *aleatorio* de selección de la *muestra* (los casos a observar). La *aleatorización* favorece la heterogeneidad en la composición de la *muestra,* como se verá en el Capítulo 5.

EJEMPLO DE VARIABLES ALEATORIAS

En el ejemplo anterior, quizás algunas de las personas con un índice elevado de alcoholemia sufriera el accidente al maniobrar para salvar un obstáculo (para no atropellar a un perro, por ejemplo). La proporción de casos (respecto del total analizado), en los que dicha circunstancia concurre, será muy baja (apenas unos casos). Por esta razón, la no consideración de la incidencia de esta variable en la investigación no introduciría sesgos importantes en los resultados de la indagación. De ahí su calificación como *variable aleatoria* en esta investigación.

• *Tipos de variables según su nivel de abstracción*

Un último criterio de clasificación hace referencia al nivel de abstracción de la variable. De acuerdo con él, existen *variables generales, intermedias* e *indicadores.*

1) *Variables generales*

Aquellas variables que son tan genéricas y abstractas que no pueden ser directamente observadas. Su medición exige que se traduzcan a *variables intermedias* e *indicadores.*

Un ejemplo típico de *variable genérica* lo representa la variable "status social" porque necesita de *indicadores* concretos que ayuden a su medición.

2) *Variables intermedias*

Expresan alguna *dimensión* o aspecto parcial de los comprendidos en la *variable genérica.*

Por ejemplo, el "nivel educativo" para la medición de la variable "status social".

3) *Indicadores* o *variables empíricas*

Representan aspectos específicos de las *dimensiones* que comprende un *concepto abstracto* o *variable genérica.* Se distinguen por ser directamente medibles.

Por ejemplo, los "cursos académicos cumplidos" como *indicador* para la *dimensión* "nivel educativo".

En la siguiente sección se desarrollará el *proceso de operacionalización de conceptos teóricos:* cómo se pasa de *variables generales* a *indicadores* e *índices*. El propósito es transformar el *concepto teórico* en observable.

4.3. De los conceptos teóricos a los indicadores e índices

Para medir la ocurrencia de *conceptos teóricos,* se procede a su *operacionalización*. Primero, proporcionando una *definición operativa,* que comprenda el significado determinado que se da al *concepto*. Segundo, especificando los *indicadores empíricos* que representarán a los *conceptos teóricos,* y se utilizarán para medir el grado de existencia del *concepto* en determinados contextos.

En la indagación científica, la delimitación de los *conceptos* depende de dos tipos de definiciones: la *definición nominal* y la *operacional*.

a) Una *definición nominal* o *teórica* es aquella que se asigna a un *concepto* (para definirlo y perfilarlo), pero que carece de las precisiones necesarias para medir los fenómenos a los que hace referencia el *concepto*.

Por *concepto* puede entenderse –siguiendo a Bollen (1989)– cualquier idea que vincule diferentes fenómenos (actitudes, comportamientos) bajo una misma etiqueta. Por ejemplo, el concepto de "alegría" reúne las distintas manifestaciones (verbales y físicas) de los individuos: reír, cantar, gritar, dar saltos, besos, abrazos.

b) La *definición operacional,* en cambio, especifica cómo se medirá la ocurrencia de un *concepto* determinado en una situación concreta. En dicha definición se detallará el contenido del *concepto* que va a medirse, vinculando los *constructos* definidos teóricamente con los *procedimientos operacionales*.

Teóricamente, la *definición operacional* consiste en "la identificación de operaciones de investigación con ayuda de las cuales se puede decidir acerca de la presencia y la intensidad de aquellos hechos que permitirán la deducción de la presencia de los fenómenos conceptualmente caracterizados" (Mayntz *et al.,* 1983: 28).

EJEMPLOS DE DEFINICIÓN TEÓRICA Y OPERACIONAL

- *Concepto:* "El absentismo (escolar)".
- *Definición teórica:* "Ausencia de la escuela en horario escolar".
- *Definición operacional:* "Número de días al mes que el alumno falta a la escuela sin justificación".

La *operacionalización de conceptos teóricos* fue pioneramente desarrollada por Lazarsfeld (1973a). Este destacado metodólogo distinguió las siguientes *fases* en el desarrollo de la *operacionalización:*

1) *Representación teórica del concepto* de forma que queden reflejados sus rasgos definitorios.
2) *Especificación del concepto,* descomponiéndolo en las distintas *dimensiones* o aspectos relevantes que engloba.
3) Para cada una de las *dimensiones* elegidas se selecciona una serie de *indicadores* (o *variables empíricas),* que "indiquen" la extensión que alcanza la *dimensión* en los casos analizados.
4) *Síntesis de los indicadores* mediante la elaboración de *índices.* A cada *indicador* se le asigna un *peso* o *valor,* de acuerdo con su importancia. A partir de estos *valores* se confecciona un *índice,* una medida común que agrupe a varios *indicadores* de una misma *dimensión conceptual* operacionalizada numéricamente.

A este proceso de *operacionalización* diseñado por Lazarsfeld, hay que hacer –de acuerdo con González Blasco (1989)– las precisiones siguientes:

a) Por muchas *dimensiones* que se consideren, nunca puede abarcarse la totalidad de un *concepto.* Sobre todo, cuando el *concepto* incluye una gran variedad de aspectos.
b) Operar con demasiadas *dimensiones* de un mismo *concepto* puede dificultar los análisis. Por esta razón resulta conveniente limitar el número de *dimensiones* a aquellas que sean más relevantes en el *concepto* considerado. Pero, como no existe ninguna regla teórica que contribuya a la delimitación del número de *dimensiones* a considerar, ésta se deja a la intuición del investigador.
c) Después de haber escogido los *indicadores,* ha de volverse a considerar el fenómeno que se observa. Con ello se pretende comprobar si las medidas obtenidas mediante los *indicadores* reflejan los hechos observados. En caso afirmativo, se estudia la necesidad de ajustar el *concepto* que los hechos reflejen.

Los *indicadores* constituyen propiedades esencialmente *manifiestas* que, supuestamente, se hallan empíricamente relacionadas con una *propiedad latente* o no observable *(dimensión).* De ellos se exige que sean "expresión numérica, cuantitativa de la dimensión que reflejan" (González Blasco, 1989: 235). Por lo que, en cualquier *operacionalización de un concepto* habrá que encontrar, al menos, un *indicador* para cada *dimensión* del *concepto.*

No obstante, ha de tenerse presente que los *indicadores* representan "aproximaciones" (en términos de "probabilidad") al *concepto* que miden. Ello se debe, pre-

cisamente, a las características de generalidad y abstracción que definen a todo *concepto*.

En consecuencia, tomar las "detenciones policiales" como único o principal *indicador* de "delincuencia", por *ejemplo,* sería inexacto. Puede haberse cometido un acto delictivo y no haberse registrado ninguna detención policial o, a la inversa, una persona puede ser detenida sin haber cometido ningún acto delictivo. De ahí la recomendación –ya reiterada en capítulos precedentes– de acudir a la *operacionalización múltiple* como recurso para aumentar la *validez* de una investigación *(validez de constructo).*

Las razones a favor de la *operacionalización múltiple* (o *medición* de un mismo *concepto* por medios diferentes) pueden resumirse en dos fundamentales:

a) Utilizar varias *medidas* para un mismo *concepto* contribuye a abarcar las distintas *dimensiones* que éste incluye.
b) Proporciona una mayor precisión y validez de la *medición,* cuando coinciden los resultados de mediciones diferentes.

La elección de *indicadores* depende de los *objetivos de la investigación,* principalmente. Además se recomienda:

a) Confeccionar una lista con el mayor número de *indicadores* posible. Después, a la vista de la información reunida, se procederá a la eliminación de aquellos *indicadores* que resulten ser no significativos para la medición de esa *dimensión* concreta.
b) Acudir a *indicadores* ya validados en investigaciones previas. Ello favorece la comparación de los resultados alcanzados en estudios diferentes, además de suponer una garantía mayor en su aplicación.

   La adecuación de un *indicador* puede variar en distintos contextos y momentos temporales. Si se recurre a aquéllos ya validados en indagaciones previas, podrá analizarse dicha variación.

El *análisis factorial* resulta de gran utilidad en la elucidación de las distintas *dimensiones* que comprende un *concepto,* pero con posterioridad a la recogida de información.

Esta *técnica analítica multivariable* se ocupa del análisis de las interrelaciones entre un gran número de variables (tomadas de respuestas a preguntas de un *cuestionario,* por ejemplo). El objetivo es explicar dichas variables en términos de sus *dimensiones latentes comunes*, denominadas *"factores"*. La condición exigida es que el número de *factores* (o de *dimensiones latentes)* sea inferior al número de variables originarias en el análisis.

Cada *factor* queda definido por las variables que muestren una mayor correlación respecto a él que hacia los otros *factores*. El estadístico más empleado para medir el

grado de *correlación* entre la *variable empírica* (o *indicador)* y el *factor* (o *variable latente)* es el *"factor loading"*. Este estadístico ha de presentar un valor elevado (usualmente superior a .30) para que la correlación entre el *indicador* y el *factor* se considere significativa. Cuanto más elevado sea el valor del *"factor loading"*, mayor será la contribución del *indicador* a la configuración del *factor* o *dimensión* del *concepto*.

Una vez definidos los *factores,* el investigador les adjunta una *"etiqueta"* (o nombre) que proporcione una denominación común a los distintos *indicadores* que representan cada *factor* (o *dimensiones del concepto* que se mide). Para ello se parte de la composición de cada *factor*; es decir, de las *variables empíricas* que lo configuran.

En la sección 4.4. se ejemplifica la utilización del *análisis factorial* en la *operacionalización de conceptos*. En el Capítulo 9 se abordará la *técnica multivariable del análisis factorial,* junto a otras técnicas estadísticas de análisis.

Finalmente, el *concepto* queda perfilado cuando se conocen los distintos valores que presenta cada una de las *dimensiones* parciales en que se divide. Los *indicadores* son los que proporcionan la información numérica. A partir de ésta se efectúa una serie de operaciones aritméticas, de las que resulta la obtención de un *índice:* una cifra resumen (y numérica) de los diversos *indicadores* de un *concepto*. Véase, más adelante, los ejemplos de la sección 4.4.

Para el *cálculo de un índice* se precisa que las distintas medidas se transformen en una *escala de medición* común, con la finalidad de facilitar su agregación. Este proceso de consecución de un *índice* suele acompañarse de la *ponderación*.

Ponderar supone asignar "pesos" a los distintos valores que presentan los *indicadores,* en un intento de expresar diferencias en la importancia relativa de cada uno de los *indicadores* en el *índice* compuesto.

Aunque la *ponderación* se realice siguiendo una lógica determinada, generalmente resulta arbitraria. En el *Informe FOESSA 1970* (dirigido por Amando de Miguel) se resumen en cuatro las recomendaciones esenciales a seguir en la elaboración de cualquier *coeficiente de ponderación* (FOESSA 1970: 1604-1605):

a) Representar lo más fidedignamente posible a la variable que se pondera y las diferencias de sus indicadores.

Esta característica se puede cumplir con cierta facilidad en algunos casos; por ejemplo, cuando la frecuencia tiene una simbología cuantitativa:

Valoración	*Coeficiente ponderación*
Diariamente	30
Casi todos los días	20
Semanalmente	4
Varias veces al mes	2
Casi nunca	1
Nunca	0

Pero, en otros casos, requiere una abstracción para ponderar de forma distinta reacciones no equivalentes unas de otras. Por ejemplo, el estímulo (en este caso) es el grado de creencia respecto de varios dogmas de la religión católica:

Valoración	*Coeficiente de ponderación*
Cree firmemente	5
Cree con algún reparo	3
Cree a medias	2
No cree en absoluto	0

La ponderación aquí efectuada se basa en la idea del propio investigador de que "creer a medias" y "creer con algún reparo" se distinguen bastante menos que "creer firmemente" (que supone una mayor distancia respecto de "no creer del todo").

b) El coeficiente ha de ser sencillo. A ser posible, un número entero y pequeño.

c) Deben utilizarse los signos (+) y (–) siempre que el resultado final del *índice* indique que es una variable con dos significaciones bien distintas. Éste es el caso, por ejemplo, en las reacciones ante el estímulo de varios ítems que representan frases-tópico, ante las cuales se puede estar "muy de acuerdo" o "muy en desacuerdo":

Valoración	*Coeficiente de ponderación*
Completamente de acuerdo	+2
Algo de acuerdo	+1
Algo en desacuerdo	–1
Completamente en desacuerdo	–2

d) Las frecuencias o atributos iguales han de ponderarse de igual forma. Ello permite la comparación posterior de los *índices*. Lo que representa una ventaja indiscutible para el análisis y la interpretación de los datos.

Por ejemplo, en la pregunta X se mide el grado de "religiosidad subjetiva" del entrevistado; en la pregunta Y, la religiosidad del "ambiente en el que se vive, sus amistades y las personas que conoce". En ambos casos, se han seguido los mismos criterios de *ponderación:*

Valoración	*Coeficiente de ponderación*	
	Pregunta X (religiosidad subjetiva)	*Pregunta Y (religiosidad grupo referencia)*
Muy buen católico	7	7
Católico practicante	5	5
Católico no muy praticante	3	3
Católico no practicante	1	1
Indiferente en religión	0	0
Creyente de otra religión	0	0

Los *coeficientes de ponderación* así obtenidos se multiplican por las frecuencias de cada ítem, en cada valoración. Para simplificar y estandarizar en diversos ítems, se utilizan las *frecuencias relativas* en forma de porcentajes, fundamentalmente. Para ello debe cumplirse el principio de "calcular los porcentajes perpendicularmente a la dirección de la comparación" (FOESSA 1970: 1604-1605).

EJEMPLOS DE PONDERACIÓN PARA EL CÁLCULO DE ÍNDICES

Como ilustracción de lo expuesto, se extractan dos tablas tomadas de sendos *Informes de la Universidad Complutense:*

a) Relación entre estado de salud y sentimiento de soledad, controlando por sexo

Pobl. 65 y más (1992) % verticales	*Frecuencia de sentimiento de soledad, según sexo*			
	Varones		*Mujeres*	
Estado de salud	*Sí**	*No***	*Sí**	*No***
Muy bueno	15	17	9	21
Bueno	37	56	36	47
Regular	35	23	41	27
Malo	13	5	12	4
Muy malo	1	0	1	1
Índice***	3,55	3,88	3,37	3,83

* "Sí" abrevia las categorías de respuesta "muchas y algunas veces".
** "No" abrevia las categorías de respuesta "nunca o casi nunca".
*** El *índice* se ha calculado ponderando cada porcentaje correspondiente al grado de salud (5, 4, 3, 2, 1) y se ha dividido por el sumatorio de porcentajes.
Fuente: Valles, M. y Cea, Mª A. (1994), "Los Mayores". En Miguel, A. de: *La Sociedad Española, 1993-1994*, Alianza, p. 865.

Por *ejemplo,* el *índice* correspondiente al "sí" de los varones se obtiene:

$$I = \frac{(15 \cdot 5) + (37 \cdot 4) + (35 \cdot 3) + (13 \cdot 2) + (1 \cdot 1)}{100} = 3{,}55$$

b) Comunicación telefónica entre padres e hijos

Pobl. 65 y más con hijos (1993) % horizontales	*¿Con qué frecuencia habla por teléfono?*					
	A diario	*Varias veces a la semana*	*Una vez a la semana*	*Una vez al mes*	*Una vez en meses*	*Indice* (días/mes)*
Varones	40	27	20	7	6	16,11
Mujeres	40	27	20	8	5	16,12
– 75 años	43	27	18	7	5	16,93
75 y más	35	26	24	8	7	14,66
Total	40	27	20	7	6	16,11

* El *índice* se ha calculado ponderando cada porcentaje por el coeficiente correspondiente a los días del mes (30, 12, 4, 1, 0) y se ha dividido por el sumatorio de porcentajes
Fuente: Valles, M. y Cea, Mª A. (1995), "Persona y Sociedad en la Vejez". En Miguel, A. de: *La Sociedad Española, 1994-1995*, Ed. Complutense, pp. 810.

Por *ejemplo,* el *índice* correspondiente a los varones sería:

$$I = \frac{(40 \cdot 30) + (27 \cdot 12) + (20 \cdot 4) + (7 \cdot 1) + (6 \cdot 0)}{100} = 16{,}11$$

Y el *índice* correspondiente a las personas de 75 años y más:

$$I = \frac{(35 \cdot 30) + (26 \cdot 12) + (24 \cdot 4) + (8 \cdot 1) + (7 \cdot 0)}{100} = 14{,}66$$

4.4. Ejemplos de aplicaciones de conceptos-indicadores-índices tomados de investigaciones reales

Para ilustrar la práctica de la *operacionalización de conceptos teóricos,* se han seleccionado dos investigaciones publicadas en el año 1993. La primera, ejemplifica la

aplicación del *análisis factorial* en la *medición de conceptos teóricos*. La segunda, la elaboración del *marco teórico* previo al desarrollo empírico de la *operacionalización*.

- *Cano, J. I.; Díaz, P.; Sánchez, A. y Valles, M. S. (1993):* El desarrollo social de los municipios zamoranos, *Instituto de Estudios Zamoranos "Florián de Ocampo", Zamora.*

Esta investigación tiene como objetivo principal conocer cómo se estructura económica, social y espacialmente la provincia de Zamora. Para ello, el equipo investigador analiza, en primer lugar, el nivel de *desarrollo social* de los distintos municipios que componen Zamora para, en segundo lugar, establecer una *tipología de municipios* que mejore la comprensión de la estructuración de dicho espacio.

Primero, se operacionaliza el *concepto* de "desarrollo social". Conscientes de que dicho *concepto* no es unívoco, sino más bien polémico, los investigadores optan por una operacionalización que destaca tres *dimensiones conceptuales* concretas:

a) *Dimensión demográfica.* Se piensa que el perfil demográfico de la población refleja (además de su propio valor) otra serie de fenómenos que exceden a la propia demografía.

b) *Dimensión de accesibilidad y comunicación.* Se centra en el componente espacial, de especial trascendencia en pequeños núcleos de población. A menudo, es el transporte más que el nivel de equipamientos del propio enclave lo que más va a condicionar su nivel de vida.

En consecuencia, se trata de medir la disponibilidad y el acceso a bienes y servicios.

c) *Dimensión socioeconómica, de bienestar y consumo.* Se pretende analizar el dinamismo económico y social de los municipios zamoranos.

Para la *medición* de cada una de estas tres *dimensiones,* se buscan los *indicadores* que muestren ser más adecuados (dentro de los límites de las fuentes de información disponibles).

Las *variables demográficas* de las que se disponía al comienzo de la investigación incluían *indicadores* de recuento correspondientes a los dos últimos momentos censales –cuando se realizó la investigación– (1970 y 1981), padronales (1976 y 1986), *índices* que reflejan la evolución de la población en esos intervalos, e *indicadores* de flujos *(movimiento natural de la población),* también relativos a esos períodos de tiempo. Más concretamente, se partió de las *variables* siguientes:

a) Tasa de jóvenes de los años 1981 y 1986.
b) Tasa de viejos de los años 1981 y 1986.
c) Tasa de dependencia de los años 1981 y 1986.
d) Tasa de masculinidad del año 1986.
e) Tasa bruta de natalidad de los años 1976-1980 y 1981-1985.

f) Saldo migratorio de los mismos años.
g) Incremento poblacional interpadronal (1975-1986).
h) Incremento poblacional intercensal (1970-1981).

A cada *indicador* se le acompaña de su definición, concretando: objetivos, objeciones, representación espacial y fuentes.

A continuación se realiza un análisis de cada uno de estos *indicadores* por separado. Posteriormente, se aplica la técnica multivariable de *análisis de componentes principales* para medir el peso e importancia de cada *indicador* en la *dimensión* considerada, además de la obtención de uno o varios *índices* (o *factores*) por cada una de las tres *dimensiones*. Los *indicadores* que mostraron ser significativos fueron los siguientes:

a) Tasa de jóvenes de 1986 (TASJO86).
b) Tasa de viejos de 1986 (TASVIE86).
c) Incremento interpadronal 1975-1986 (INCRPAD).
d) Incremento intercensal 1970-1981 (INCRCEN).
e) Tasa bruta de natalidad de 1981 (TBN81).
f) Tasa bruta de mortalidad de 1981 (TBM81).
g) Saldo vegetativo medio del período 1981-1986 (SALVEG81).
h) Saldo migratorio medio del período 1981-1986 (SALMIG81).

El *análisis de componentes principales* realizado con estas variables define dos *factores* con las saturaciones siguientes *("factor loadings")*.

	Factor 1	Factor 2
Tasjo86	.84051	.05365
Tasvie86	–.75027	.07639
Incrpad	.76192	.47921
Incren	.63833	.01963
Tbn81	.73888	–.00758
Tbm81	–.62946	.54007
Salveg81	.86297	–.36133
Salmig81	.35076	.83228

Fuente: Cano *et al.* (1993: 25).

Estas saturaciones pueden interpretarse como el peso de cada variable en la formación del *factor*. El signo indica si el *indicador* es directa (signo positivo) o inversamente (signo negativo) proporcional a la puntuación del *factor*. La puntuación del *factor* se obtiene multiplicando cada saturación por su indicador correspondiente estandarizado (res-

tándole su *media* y dividiéndole por su *desviación típica* para que todos se hallen en la misma unidad de medida) y sumando todos estos términos.

En este caso, el *primer factor* explica el 50,8% de la varianza total y el *segundo factor* el 16,9%. Al *primer factor* se le denomina "factor de pujanza demográfica", por el elevado porcentaje de varianza explicada, y la fuerte saturación que en él presentan variables cruciales para el crecimiento demográfico (saldo vegetativo e incrementos poblacionales). El *segundo factor* responde más al "saldo migratorio" y, en menor medida, a la tasa de mortalidad y al incremento interpadronal.

Asimismo se procede con las otras dos *dimensiones* diferenciadas del *concepto* de "desarrollo social". A partir de toda la información analizada, se elabora un *índice* de orden superior, en el que se sintetizan los *indicadores* que han obtenido un mayor poder analítico en las tres *dimensiones* (un *índice* sintético del "desarrollo social"). En la Figura 4.1 se detallan los distintos *indicadores* y el *índice* obtenido.

Concepto	*Dimensiones*	*Indicadores*	*Índices*	*Índice general*
Desarrollo social	Demográfica	Tasa jóvenes Tasa viejos Incremento interpadronal Incremento intercensal Tasa bruta natalidad	Pujanza demográfica	Índice de desarrollo social
		Tasa bruta mortalidad Saldo vegetativo Saldo migratorio	Recesión demográfica	
	De accesibilidad	Distancia a la capital Dist. núcleo > 5.000 h. Nº autobuses y trenes Nº vehículos privados Dist. centros educativos	Comunicación/ accesibilidad	
	Socioeconómica	Tierras cultivadas Caballos potencia tractores por hectárea cultivada. % Tierras en barbecho.	Calidad de la tierra	
		Tasa de paro Consumo electricidad de alta tensión por habitante. % Pobl. ocupada en sector secundario.	Industrialización	
		Nº medio teléfonos por hab. Consumo electricidad de baja tensión por habitante. % Pobl. con estudios de secundaria.	Bienestar y terciarización	

Figura 4.1. Indicadores e índices del concepto de "desarrollo social" de los municipios de Zamora. Basado en Cano *et al.* (1993).

Después de elaborar medidas del "desarrollo social" de forma sectorial, se compone un *índice único*, que resuma el grado de "desarrollo social". Los *factores (índices)* obtenidos tienen una buena parte en común (atendiendo a su *matriz de intercorrelaciones* que proporciona el *análisis factorial de componentes principales).* Si bien, cada uno mide una cosa parcialmente distinta.

Se realiza un nuevo *análisis de componentes principales* con los *factores*. De él resulta un *superfactor*, que mide el "desarrollo social" (y explica el 56,7% de la varianza total). Este *metafactor* está integrado, principalmente, por los cuatro *factores* siguientes: pujanza demográfica (FACDEM); de comunicaciones (FACOM); de bienestar y servicios (NFACSERV); y de productividad, rendimiento y mecanización de la agricultura (FACALTER), abreviado con la expresión "calidad de la tierra".

El *factor* que menos correlaciona con el resto es el de "calidad de la tierra". Lo cual no es de extrañar, dado que es el único *factor* que no se aplica a todos los municipios por igual, sino sobre todo a aquéllos en los que la agricultura tiene un peso importante.

Índice general	
Factor 1	
Facdem	.82034
Facom	.78623
Nfacserv	.78765
Facalter	.59841

Fuente: Cano *et al.* (1993: 140).

Estos valores expresan la contribución de cada *factor* a la elaboración del *superfactor*. Pero, a diferencia de lo realizado en la creación de los *factores* originales (o *índices),* en la confección del *índice general* los autores no se limitan a estandarizar los componentes y multiplicarlos por sus saturaciones; sino que deciden, además, multiplicar cada *factor estandarizado* por una *ponderación* elegida, según la importancia teórica que conceden a cada uno de los *factores*. Las *ponderaciones* que se han utilizado son las siguientes:

Ponderaciones de los factores (índices)	
Facdem	* 6
Facom	* 3
Nfacserv	* 3
Facalter	* 2

Fuente: Cano *et al.* (1993: 140).

Por último, se introduce una lista con las puntuaciones concretas de cada uno de los municipios de Zamora en este *índice de desarrollo social.* Lo que interesa es conocer la distancia de cada municipio zamorano en términos de "desarrollo social".

- *Setién, M.ª Luisa (1993):* Indicadores sociales de calidad de vida. Un sistema de medición aplicado al País Vasco. *Madrid, CIS, Monografía n.º 133.*

Esta segunda investigación tiene como objetivo fundamental la elaboración de un sistema de *indicadores sociales* para medir la "calidad de vida".

El *concepto* "calidad de vida" constituye un concepto también abstracto, complejo e indirectamente medible. Por tanto admite múltiples orientaciones y definiciones variadas. Como consecuencia, el primer problema que se presenta consiste en la elaboración de un *marco conceptual,* que delimite los significados, además de proporcionar un marco coherente al conjunto de los *indicadores sociales* que se deduzcan.

Concibiendo la "calidad de vida" como un concepto vinculado al "desarrollo", su estudio se enmarca en una teoría del desarrollo. Concretamente, la autora parte de tres teorizaciones principales:

a) La teoría del desarrollo de Galtung y Wirak (1979).
b) La teoría de las necesidades de Maslow (1981).
c) La teoría de las discrepancias múltiples de Michalos (1985).

A partir de estas teorías, se dan las definiciones siguientes a cada uno de los *conceptos* esenciales en la investigación:

a) *Desarrollo:* "proceso de realización de las potencialidades humanas mediante la satisfacción de las necesidades, utilizando para ello los medios societales y contando con un entorno natural limitado".
b) *Necesidades:* "aquello que es menester para la vida humana, sea física, mental, espiritual o social".
c) *Calidad de vida:* "grado en que una sociedad posibilita la satisfacción de las necesidades (materiales y no materiales) de los miembros que la componen".

De la satisfacción de las necesidades dependerá el grado de calidad con que puede calificarse la vida social en los *continua* peor-mejor, bueno-malo, en que puede situarse. De esta forma, el *marco conceptual* seguido en la investigación orienta la medida de la "calidad de vida" a la satisfacción de las necesidades.

En suma, son las necesidades las que proporcionan la base para encarar el sistema de *indicadores sociales*. De lo que se trata es de establecer unas áreas de necesidad o componentes de la "calidad de vida", que representen el amplio abanico de necesidades. Para su *operacionalización* se opta por:

a) Elaborar amplias listas de necesidades (físicas, psicológicas, espirituales, sociales, culturales), siguiendo el criterio de la experiencia.
b) Analizar comparativamente los sistemas de *indicadores sociales* que ya se han establecido con anterioridad, siguiendo el criterio del "consenso político".

De la conjunción de ambos criterios resulta la selección de 11 áreas o componentes de "calidad de vida", que parecen abarcar todos los ámbitos de las necesidades humanas:

1) Salud.
2) Renta (aspecto material-bienes).
3) Trabajo.
4) Vivienda.
5) Seguridad.
6) Educación-Formación.
7) Familia.
8) Entorno físico-social (medio ambiente físico y medio ambiente social que enmarca la vida cotidiana –relaciones interpersonales, vecinales y de amistad– integración en grupos, organizaciones, en la comunidad).
9) Ocio-tiempo libre.
10) Religión.
11) Política (participación, libertad, prestigio e identidad comunitaria).

Después, se determinan las *dimensiones* que comprende cada área que compone la "calidad de vida". A partir de estas *dimensiones* se seleccionan los *indicadores* que mejor midan cada una de las *dimensiones*.

La *dimensión* se define como una categoría genérica en la que se divide el área, estando referida a algún aspecto importante de la misma. En total se diferenciaron 39 *dimensiones* (algunas de ellas incluso divididas, a su vez, en *subdimensiones)* y 251 *indicadores,* que dan contenido al sistema para la medida de la "calidad de vida". Como ejemplo, en el *área* de salud se distinguieron las cuatro *dimensiones* siguientes:

1) Estado de salud de la población y su distribución.
2) Atentados contra la salud.
3) Educación sanitaria.
4) Recursos para la salud.

En la *dimensión* primera, el "estado de salud", se pretende medir la duración de la vida de la población en la comunidad. Para ello se utilizan varios *indicadores:*

1) Esperanza de vida (al nacer y a la edad de 1, 20, 40 y 60 años).
2*a)* Tasa de mortalidad.
2*b)* Tasa de mortalidad específica por edades.
3) Tasa de mortalidad según la causa de muerte.
4) Mortalidad perinatal.

Para mostrar la "salud durante la vida", se seleccionan los siguientes *indicadores:*

5) Morbilidad hospitalaria.
6) Autovaloración del estado de salud.
7) Incapacidad permanente.
8) Incapacidad transitoria.
9) Padecimientos físico/psíquicos.
10) Distribución del estado de salud.

Para la segunda *dimensión* ("atentados contra la salud") se distinguen seis *indicadores:*

11) Relación peso/talla de los niños.
12) Consumo de tabaco.
13) Consumo de alcohol.
14) Hábitos de ejercicio físico.
15) Nivel de arraigo de modo de vida atentatorios contra la salud.
16) Conocimiento sobre los atentados contra la salud.

Sobre las demás *dimensiones* del *concepto* de "calidad de vida" remito al lector interesado a la citada obra (editada en la colección *Monografías* del CIS, con el n.º 133), dado el voluminoso número de *indicadores* que la autora destaca para cada *dimensión.*

Por último, conviene señalar que esta segunda investigación aquí comentada se limita al *desarrollo conceptual* de la medida de la "calidad de vida"; es decir, se centra en la traducción de nociones abstractas a términos concretos y cuantificables, además de en la especificación de las *medidas* (o *indicadores)* pertinentes.

Sin embargo, este estudio, a diferencia del anteriormente referido, no trata la otra cara de la *medición:* el *desarrollo de datos* (el sistema de recogida y tratamiento de la información). Se queda, por tanto, en el nivel teórico-metodológico, no extendiéndose al metodológico-técnico.

4.5. Cuestiones de validez y de fiabilidad en la medición

Una vez que se han seleccionado los *indicadores,* el siguiente paso será comprobar hasta qué punto la operacionalización de los *conceptos teóricos* realizada reúne unas condiciones mínimas de *validez* y de *fiabilidad.*

4.5.1. La validez de la medición

Antes que *fiables,* los *indicadores* han de ser *válidos.* Es decir, han de proporcionar una representación adecuada del *concepto teórico* que miden. Ello es independiente de si reúne o no condiciones de *fiabilidad.*

La *validez* hace referencia a la relación que ha de existir entre el *concepto teórico* y el *indicador empírico.* Concretamente, el investigador ha de comprobar si los *indicadores* elegidos realmente "indican" lo que se pretende que indiquen; si "miden" correctamente el *significado* dado al *concepto teórico* en consideración.

Mediante esta comprobación el investigador tratará, en suma, de identificar si las *definiciones operacionales* e *indicadores* aplicados resultan apropiados para la medición del *concepto* de interés.

Si, por *ejemplo,* se emplea la variable "nivel de ingresos" como un *indicador* de "status social", habría que comprobar –de acuerdo con el criterio de *validez*– si el "nivel de ingresos" puede considerarse un *indicador* "válido" de la variable "status social".

La *validez de la medición* depende de cómo se haya definido y operacionalizado el *concepto* que se analiza. Carmines y Zeller (1979) diferencian tres modalidades básicas de *validez:*

a) Validez de criterio.
b) Validez de contenido.
c) Validez de constructo.

• *Validez de criterio*

A veces también referida como *validez predictiva* o *concurrente.* En ella, la *validez de la medición* se comprueba comparándola con algún "criterio" (o *medida* generalmente aceptada) que anteriormente se haya empleado para medir el mismo *concepto.* El objetivo es demostrar que la nueva *medida* clasifica a los individuos, u otras unidades de análisis, de igual forma que otros *indicadores* alternativos de la misma *variable latente.*

En general, la utilización de varios criterios contribuye a aumentar la confianza en la *medición.*

EJEMPLO DE VALIDEZ DE CRITERIO

Comparar datos sociodemográficos obtenidos mediante *encuesta* con los registrados en el último *censo de población* o *padrón de habitantes.* Si los datos coinciden, éstos se consideran válidos.

Técnicamente se diferencian dos variedades de *validez de criterio:* la *validez concurrente* y la *predictiva.*

a) *Validez concurrente.* Cuando se correlaciona la medición nueva con un criterio adoptado en un mismo momento.
   Por ejemplo, comparar los datos de *encuesta* sobre intención de voto con los datos oficiales tras celebrarse las elecciones.
b) *Validez predictiva.* Concierne a un criterio futuro que esté correlacionado con la *medida.*
   Por ejemplo, comparar las respuestas dadas en una *encuesta* sobre racismo, realizada a empresarios, con la conducta que posteriormente éstos manifiestan en la contratación de empleados.

Este procedimiento de comprobación de la *validez* de una *medición* no se halla, sin embargo, carente de polémica. De Vaus (1990) apunta dos *problemas* principales:

a) Se asume la *validez de la medida* establecida con preferencia a la nueva. Si se observa una baja correlación entre la medida nueva y la establecida, implica que la *medida* nueva se toma como "inválida". Pero, tal vez la inválida sea la antigua (la tradicionalmente aceptada), y no la nueva *medida.*
b) Muchos *conceptos* en las ciencias sociales carecen de mediciones (generalmente aceptadas) que puedan emplearse para la comprobación de nuevas mediciones. Dicha probabilidad aumenta, cuanto más abstracto es el *concepto* que se analiza.

• *Validez de contenido*

Concierne al grado en que una *medición* empírica cubre la variedad de significados incluidos en un *concepto.*

EJEMPLO DE VALIDEZ DE CONTENIDO

Si el *concepto* que se mide es el de "racismo", habría que considerar –siguiendo el *criterio de validez de contenido*– si se han comprendido las diversas manifestaciones de racismo o si, por el contrario, la *medición* se ha circunscrito a unos pocos aspectos y hechos concretos.

El recurso a *múltiples indicadores* contribuye a reducir este tipo de *invalidez.*

• *Validez de constructo*

Cuando se compara una medida particular con aquella que teóricamente habría de esperar (a partir de las hipótesis derivadas del *marco teórico* de la investigación).

EJEMPLO DE VALIDEZ DE CONSTRUCTO

Si la "delincuencia" se mide exclusivamente con *datos policiales,* es probable que éstos corroboren las teorías que enfatizan la mayor comisión de actos delictivos por las clases sociales desfavorecidas.

En cambio, si la "delincuencia" se mide mediante *encuestas de autodenuncia,* la influencia de la clase social en la causalidad de la delincuencia probablemente se debilitará. De ahí que deba precisarse qué *medida* de "delincuencia" resulta más válida. La teoría de referencia puede no ser cierta y, como consecuencia, los *indicadores* de la variable clase social empleados resultan inválidos en la *medición* de la "delincuencia".

Todas estas modalidades de *validez* cabe considerarlas variedades de lo que Campbell y Fiske (1959) llamaron *validez convergente:* demostrar que una forma concreta de *medir* un *concepto* "converge" con otras maneras distintas de medirlo. Estos autores propusieron –como ya se comentó en el Capítulo 2– la utilización de varias técnicas de obtención de información, como táctica para validar la *medición* hecha mediante otra técnica (o instrumento de *medición).*

Además de la *validez convergente,* Campbell y Fiske (1959) diferencian la *validez discriminante* o *divergente:* una medida debería mostrar bajos niveles de correlación con medidas de *conceptos* diferentes. E insisten, nuevamente, en la necesidad de recurrir a distintas técnicas de obtención de información para garantizar, también, esta segunda modalidad de *validez.*

En suma, ambos tipos de validez se asegurarán midiendo el mismo *concepto* de muchas formas: cuanta mayor diversidad presenten los datos, mayor seguridad habrá en la *validez* de sus resultados. Por esta razón se invita a seguir la siguiente recomendación:

> "Si no existe ningún acuerdo claro en cómo medir un concepto, mídelo de varias formas distintas. Si el concepto tiene diversas dimensiones diferentes, mídelas todas. Y, por encima de todo, tienes que saber que el concepto no tiene ningún significado distinto del que le dimos" (Babbie, 1992: 134).

4.5.2. La fiabilidad de la medición

La *fiabilidad* se refiere a "la capacidad de obtener resultados consistentes en mediciones sucesivas del mismo fenómeno" (Jacob, 1994: 363). Quiere esto decir, que los

resultados logrados en mediciones repetidas (del mismo *concepto)* han de ser iguales para que la *medición* se estime *fiable*.

Una forma común de comprobar la *fiabilidad* consiste en aplicar el mismo procedimiento de *medición* en diferentes momentos para, posteriormente, observar si se obtienen resultados similares en las distintas *mediciones* del *concepto*. Pero, existen además otros métodos de comprobar la *fiabilidad*. Éstos pueden resumirse en cuatro métodos referidos a dos conceptos incluidos en el de *fiabilidad: estabilidad* y *consistencia*.

a) Método test-retest.
b) Método alternativo.
c) Método de las dos mitades.
d) Método de consistencia interna alpha de Cronbach.

• *Método test-retest*

Constituye la forma más sencilla de comprobar la *fiabilidad*. Consiste en administrar una misma *medida* a una misma población en dos períodos de tiempo diferentes para, así, poder observar si existe variación en las respuestas.

A la misma *muestra* de individuos se le preguntarían las mismas cuestiones en dos momentos sucesivos (en el período de dos meses, por ejemplo). A continuación, se calcularían los coeficientes de correlación entre las respuestas dadas en las dos ocasiones temporales en que acontece la recogida de información. De esta manera se mide uno de los conceptos incluidos en la *fiabilidad:* la *estabilidad*.

Una *medición* será *estable* siempre y cuando se obtenga un elevado *coeficiente de correlación* entre los datos obtenidos en los dos períodos de tiempo. En cambio, si la correlación resulta baja, esto puede llevar consigo la *inestabilidad* de la *medición*.

$$\text{Fiabilidad } = 1 - \frac{\text{Diferencias observadas}}{\text{Máximas diferencias posibles}}$$

Si se consiguen exactamente los mismos resultados (en las dos administraciones de la prueba de *medición),* el *coeficiente de fiabilidad* obtenido será 1.00. Lo que significa que la *fiabilidad* se considera "perfecta": la diferencia de los datos es nula.

En general, si el *coeficiente de correlación* entre las respuestas dadas en las dos ocasiones supera o iguala el valor 0.8, se asume que dicha pregunta o *indicador* es *fiable*. En cambio, si el valor del *coeficiente de correlación* se aproxima a 0.0, la *fiabilidad* será nula.

A estas afirmaciones hay que hacer, sin embargo, las siguientes matizaciones (siguiendo a Carmines y Zeller, 1979; De Vaus, 1990):

a) Un *coeficiente de correlación* bajo no supone, necesariamente, que la *fiabilidad* sea baja. Puede significar que tras el paso del tiempo, el individuo haya cambiado en su consideración del *concepto* que se mide.

Por *ejemplo,* la actitud de una persona hacia la drogadicción puede cambiar si en ese período de tiempo conoce a personas que padezcan la drogadicción.

La probabilidad de cambios en la persona aumenta, conforme es mayor el lapso de tiempo que separa ambas mediciones.

b) También puede estar afectada por la *reactividad.* La *medición* de un fenómeno puede, a veces, inducir a cambios en el mismo fenómeno.

El individuo puede sensibilizarse hacia la cuestión que se investiga (después de haberse producido la primera *medición)* y cambiar, en consecuencia, su respuesta en una segunda *medición.* Este cambio se atribuiría a la *reactividad* generada por la primera *medición.*

c) El problema más típico de sobreestimación de la *fiabilidad* se debe a la *memoria.* El individuo puede recordar las respuestas dadas en la primera *medición* y responder, en la segunda *medición,* de la misma manera, con la intención de mostrar consistencia en sus respuestas. Esto puede inflar la *fiabilidad* aparente de la *medición.*

Para obviar este problema, se recomienda ampliar el período de tiempo comprendido entre ambas *mediciones.* Si bien, el lapso de tiempo no debe ser, tampoco, demasiado amplio. Ello dificultaría la capacidad de memoria del sujeto, pero aumentaría, a su vez, la probabilidad de producirse cambios verdaderos en la persona. Lo que contribuiría a una modificación en sus respuestas y la consiguiente subestimación de la *fiabilidad* del instrumento de *medición.*

A estas matizaciones se suma la dificultad de aplicar, de forma reiterada, un mismo instrumento de *medición* (un *cuestionario,* por ejemplo) a una misma población. Para obviar este problema, pueden hacerse las mismas preguntas (tomadas como *indicadores)* a una *muestra* más pequeña, pero de características similares a la utilizada en la investigación (De Vaus, 1990). De esta manera se reduciría la principal limitación del *método test-retest:* la experiencia que el sujeto adquiere en la primera *medición* puede influir en las respuestas que emita en la segunda *medición.*

Este método de comprobar la *fiabilidad* se desaconseja en la *medición* de fenómenos inestables (sujetos a cambio), como la intención de voto, por ejemplo.

• *Método alternativo*

Este método se asemeja al anterior. Supone analizar una misma población en momentos diferentes. Pero, difiere en un aspecto fundamental: el instrumento de *medición* varía en la segunda comprobación.

La segunda *medición* consiste en una prueba alternativa a la primera: se mide el mismo *concepto,* aunque de distinta manera. De esta forma se reduce la probabilidad del efecto de la memoria en la inflación de la estimación de la *fiabilidad.*

En su contra está el hecho –como indican Carmines y Zeller (1979)– de que este método no permite (al igual que el anterior), la distinción entre el "cambio" verdadero de la "infiabilidad" de la medición. A ello se añade la dificultad de elaborar formas alternativas de una misma *medición.*

En general, se aconseja dividir (por la mitad) un amplio grupo de *ítems,* de forma aleatoria, para realizar dos comprobaciones paralelas y poder así comparar sus *coeficientes de correlación.*

En este segundo método, como en el primero, la estimación de la *fiabilidad* se realiza comparando las correlaciones entre ambas *mediciones.*

• *Método de las dos mitades*

Representa una manera de comprobar la "consistencia" interna de una *medida.* Adquiere mayor importancia cuando se aplican múltiples *ítems* (o preguntas) para comprobar si convergen o no en la configuración de una misma *dimensión.*

A diferencia de los dos métodos anteriores, no se efectúan dos comprobaciones en períodos diferentes de tiempo, sino al mismo tiempo. Para ello se divide la serie total de *ítems* en dos mitades, y se correlacionan las puntuaciones obtenidas en la *medición.*

Este es el caso, por *ejemplo,* cuando quiere establecerse la *fiabilidad* de un *índice* integrado por 18 *indicadores.* El *índice* se administraría a un grupo y, posteriormente, se dividirían aleatoriamente los 18 *ítems* en dos grupos de 9. A continuación, se calcularía la *correlación* entre ambos grupos de *ítems.* Para ello se puede proceder de la manera siguiente: primero, se calculan todos los *coeficientes de correlación* entre los 18 *indicadores;* y, después, se halla el coeficiente promedio de todos ellos, en ambos grupos de *ítems* (Bryman, 1995).

• *Método de consistencia interna alpha de Cronbach*

Este último método es uno de los más utilizados por los investigadores sociales en los últimos años. Se obtiene calculando el promedio de todos los *coeficientes de correlación* posibles de las *dos mitades.* De esta manera se mide la *consistencia interna* de todos los *ítems*, global e individualmente.

El *alpha de Cronbach* puede calcularse a partir de la *matriz de varianza-covarianza* obtenida de los valores de los *ítems.* La *diagonal* de la *matriz* contiene la *varianza* de cada *ítem;* el resto de la *matriz* comprende las *covarianzas* entre los pares de *ítems.* Puede expresarse de la forma siguiente:

$$a = \frac{N}{N-1}\left[1 - \frac{\Sigma \text{ diagonal de la matriz}}{\Sigma \text{ todos los elementos de la matriz}}\right]$$

El coeficiente así obtenido varía de .00 *(infiabilidad)* a 1.00 *(fiabilidad* perfecta). Por lo general, no debería ser inferior a .80 para que el instrumento de *medición* pudiera considerarse "fiable".

El comando RELIABILITY del paquete estadístico SPSS facilita el cálculo del *coeficiente alpha*. Proporciona tanto los datos brutos como los estandarizados. Además, permite eliminar los *ítems* (o *indicadores)* que muestren no estar correlacionados con la puntuación global *(índice)*.

• Por último, conviene insistir en que la *infiabilidad* puede deberse a diferentes *fuentes de error:* desde la pertinencia del instrumento de *medición* (la redacción de la pregunta, por ejemplo), hasta cómo se aplicó (la actuación y características personales de los entrevistadores), y analizó la información obtenida (la codificación y el tratamiento estadístico de los datos) –como se verá en los capítulos siguientes. Lo cierto es que la *fiabilidad* perfecta resulta difícil de alcanzar. Como Carmines y Zeller (1979: 11, 12) afirman:

> "La medición de cualquier fenómeno siempre contiene una cierta cantidad de error casual (...). Porque las mediciones repetidas nunca igualan exactamente unas a otras, la infiabilidad siempre está presente, al menos a una extensión limitada."

Lecturas complementarias

Babbie, E. (1992): *The practice of social research.* Belmont, California, Wadsworth Publishing Company.

Casas, F. (1989): *Técnicas de investigación social: los indicadores sociales y psicosociales (teoría y práctica).* Barcelona, PPU.

Carmines, E. G. y Zeller, R. A. (1979): *Reliability and validity assessment.* Beverly Hills, Sage.

González Blasco, P. (1994) "Medir en las ciencias sociales". En García Ferrando, M. *et al.* (comps.): *El análisis de la realidad social. Métodos y técnicas de investigación.* Madrid, Alianza, pp. 275-334.

Latiesa, M. (1994): "Validez y fiabilidad de las observaciones sociológicas". En García Ferrando, M. *et al.* (comps.): *El análisis de la realidad social. Métodos y técnicas de investigación.* Madrid, Alianza, pp. 335-364.

De Vaus, D. A. (1990): *Surveys in social research.* London, Unwin Hyman.

Ejercicios Propuestos

1. Para cada modalidad de variable, da un ejemplo y señala los atributos que la componen.
2. Define teórica y operacionalmente los conceptos "xenofobia" y "corrupción".
3. ¿Qué dimensiones diferenciarías en la medición del nivel de "desarrollo" de un país? Para cada dimensión señala, al menos, cinco indicadores.
4. ¿En qué difiere la validez concurrente de la convergente? Añade un ejemplo a la respuesta.
5. ¿Qué se entiende por *reactividad?* Especifica distintas actuaciones dirigidas a su control.
6. Calcula un índice que resuma la información incluida en la siguiente tabla:

	¿Cómo calificaría su estado de salud?					
	Muy bueno	Bueno	Regular	Malo	Muy malo	Índice
Varones Mujeres	16 16	51 42	26 33	6 7	1 2	
– 75 años 75 y más	15 18	42 49	35 27	7 5	1 1	
Total	16	46	30	7	1	

5

LA SELECCIÓN DE LAS UNIDADES DE OBSERVACIÓN: EL DISEÑO DE LA MUESTRA

La selección de las *unidades de observación* constituye un estadio primordial en todo proceso investigador. De cómo se materialice dicha elección dependerá, en buena medida, la calidad de la información que se recoja. Razón por la cual se recomienda al investigador que no escatime, ni tiempo ni esfuerzo, en la planificación y ejecución del *diseño de la muestra.*

Pero, ¿en qué consiste el diseño muestral?; ¿cómo hay que proceder para la consecución de una *muestra* apropiada?; ¿qué es una *muestra?;* ¿qué variedades de *diseños muestrales* existen? Los apartados siguientes tratarán de dar respuesta a éstas y otras cuestiones relacionadas con las *unidades de observación* en el proceso de investigación.

5.1. Fundamentos y clarificación terminológica

Una de las primeras decisiones a tomar en cualquier investigación es la *especificación y acotación de la población* a analizar. La concreción de ésta vendrá determinada por cuál sea el *problema* y los *objetivos* principales de la investigación.

Por *población* (o *universo* de estudio) comúnmente se entiende "un conjunto de unidades, para las que se desea obtener cierta información" (Sánchez Crespo, 1971: 11). Las *unidades* pueden ser personas, familias, viviendas, escuelas, organizaciones, artículos de prensa...; y la *información* vendrá dictada por los *objetivos* de la investigación (condiciones de habitabilidad de la vivienda, reacción familiar ante el fracaso escolar, la práctica de la religión en las escuelas).

En la *definición y acotación de la población* han de mencionarse características esenciales que la ubiquen en un espacio y tiempo concreto.

EJEMPLOS DE DEFINICIÓN DE LA POBLACIÓN (O UNIVERSO) DE ESTUDIO

a) En una investigación sobre la ocupación del tiempo tras la jubilación, una posible definición del *universo* de estudio sería la siguiente:

"Población de 65 y más años que residen en el municipio de Madrid".

Además, habría que precisar si se analizará a toda la *población* (indistintamente del lugar donde resida) o si, por el contrario, se excluirá a aquellos que, en el momento de la investigación, se hallen viviendo en instituciones (como residencias para la tercera edad). Depende que cuáles sean los objetivos del estudio.

b) Si el objetivo fundamental de la investigación fuese analizar la corrupción política a través de la prensa escrita, una definición de la *población* sería:

"Artículos de prensa publicados en los periódicos *El País, ABC* y *El Mundo* entre 1990 y 1995".

Una vez definida la *población*, se procede al *diseño de la muestra:* la selección de unas *unidades* concretas de dicha *población*.

Aunque el *universo* fuese de pequeña dimensión, por razones de economía (en tiempo y dinero), rara vez se observa a cada una de las *unidades* que lo forman. Por el contrario, se decide la extracción de una *muestra* de entre los integrantes del *universo*. Si bien, en este hacer también existen divergencias. Depende, fundamentalmente, de la *estrategia de investigación* que se haya escogido para la consecución de los objetivos del estudio.

Un *estudio de casos* o un *experimento,* por ejemplo, imponen menos exigencias en la *muestra* que una *encuesta* (más motivada a la consecución de una *muestra* "representativa"). Dicha *representatividad* estará subordinada al *tamaño de la muestra* y al procedimiento seguido para la selección de las *unidades muestrales (si éste garantiza la inclusión de toda la variedad de componentes del* universo).

Si, a partir de los datos obtenidos en una *muestra*, quieren inferirse las características correspondientes de la *población (parámetros poblacionales),* es imperativo diseñar una *muestra* que constituya una representación a pequeña escala de la *población* a la que pertenece. Para ello resultan esenciales las aportaciones llevadas a cabo en el campo de la estadística. Concretamente, en la Estadística Inferencial, configurada a partir de la teoría de las probabilidades. En ella se fundamentan los *diseños muestrales probabilísticos,* como después se verá.

Cualquier *diseño muestral* (máxime los *probabilísticos)* comienza con la *búsqueda de documentación* que ayude a la identificación de la *población* de estudio. Con el término *marco* se hace referencia al "listado" que comprende las unidades de la *población*. Puede ser un censo general de la población, un registro de individuos, o cualquier otro procedimiento que lleve a la identificación de los miembros de una *población*. De

él se espera que sea un descriptor válido de dicha *población*. Por lo que ha de cumplir una serie de *requisitos* mínimos:

a) El *marco* ha de ser lo más completo posible, en orden a facilitar la representatividad de la *muestra*.

Téngase presente que la *muestra* escogida sólo podrá considerarse "representativa" de la *población* comprendida en el *marco de muestreo* elegido; es decir, de aquellos que han tenido probabilidad de participar en la *muestra*. Por esta razón, la *comprehensividad* se convierte en una exigencia básica de todo *marco muestral*.

EJEMPLOS DE COMPREHENSIVIDAD DEL MARCO DE MUESTREO

a) La guía telefónica constituye el *marco de muestreo* habitual en una *encuesta telefónica*. Pero, su uso impone una serie de restricciones en la *muestra* que el investigador ha de conocer para la acotación de la *población* a analizar.

En primer lugar, la guía telefónica (aunque esté actualizada y completa) limita la *población* a las personas con un número de teléfono "registrado". Por tanto, excluye a aquellos que carecen de teléfono. A éstos se suman los que, aun teniéndolo, no figuran en el listín telefónico (ya por expreso deseo del abonador de no figurar, ya por haberse suscrito en fechas posteriores a la última actualización de la guía, o por otra razón).

En segundo lugar, las guías no registran a todos los integrantes de la vivienda con teléfono. Sólo aparece uno de ellos; generalmente, el cabeza de familia. Este hecho puede llevar a una mayor presencia de varones en la *muestra,* si ésta se limita a las personas cuyo nombre figura en la guía telefónica.

También, se ha de tener presente que las personas con más de un número de teléfono tienen una probabilidad mayor de ser elegidas para formar parte de la *muestra*.

Todas estas circunstancias condicionan las posibilidades de generalización de la información obtenida en una *encuesta telefónica*. Esta únicamente podrá generalizarse a la *población* registrada en el *marco de muestreo* empleado; no al conjunto de la *población* que resida en el espacio geográfico que abarque el estudio.

b) Si la finalidad de la investigación fuese conocer la opinión de los sociólogos españoles sobre su actividad profesional, un *marco de muestreo* idóneo sería el directorio de doctores y licenciados en sociología colegiados. Pero su uso también impone restricciones importantes en la *población* de estudio. La opinión que se recabe sólo podría ser "representativa" de los sociólogos que estén colegiados (y no de la totalidad de los sociólogos españoles). Por lo que, habría que comprobar qué proporción representan los colegiados en el conjunto de la población.

La "representatividad" de la *muestra* también se verá afectada (aun limitándose a la *población* colegiada) por la actualización del *marco de muestreo*. Si éste no incluye a los sociólogos que se hayan colegiado en fechas recientes, y, en contra, no

excluye a aquellos que han dejado de estar colegiados (por fallecimiento u otra causa), la *muestra* podrá incluso no ser "representativa" del colectivo de sociólogos colegiados. Ello dependerá de la magnitud de los excluidos y de los incluidos erróneamente.

Si en esta segunda investigación el investigador desease una mayor "generalización" de los resultados de su indagación, tendría que recurrir a otros *marcos muestrales:* un listado de profesores de sociología, un registro de sociólogos que hayan concursado en oposiciones para las administraciones públicas, por *ejemplo.* Estos *marcos muestrales* complementarían la información proporcionada por el *marco de muestreo* principal.

En suma, cualquier *marco de muestreo* se halla limitado a un conjunto de la *población,* siendo además vulnerable a la presencia de omisiones. Del volumen de éstas, junto a la capacidad del investigador de detectarlas y, en su caso, eliminarlas, dependerá la incidencia del *marco de muestreo* en la aparición de *sesgos* en la *selección de la muestra.*

b) La *comprehensividad* del *marco muestral* conlleva, necesariamente, la exigencia de su *actualización* –como ya se ha hecho mención–.

En la medida en que el *marco muestral* se halle actualizado (en cortos períodos de tiempo, preferentemente) las posibilidades de omisiones se restringen. Por el contrario, aumenta la probabilidad de que éste contenga a los miembros reales de la población que "representa".

EJEMPLO DE ACTUALIZACIÓN DE UN MARCO DE MUESTREO

Si en 1989 se pasase una *encuesta* a la población residente en Madrid, el *marco muestral* preferible sería el *Padrón de habitantes de 1986.* Este reflejará mejor el perfil de la población madrileña –a fecha de 1989– que el último *censo de población* (que a dicha fecha sería el de 1981).

---

c) Cuando la investigación persigue la *generalización de los datos muestrales* (a la *población* que conforma el *marco muestral),* es preciso que cada componente de la población esté igualmente representado en el *marco de muestreo.*

Han de evitarse las *duplicidades.* Éste es un problema habitual cuando, para una misma *muestra,* se combinan dos o más listados diferentes.

Aquéllos que (por cualquier motivo) aparezcan más de una vez, tendrán una oportunidad superior de ser elegidos. Lo que favorece su sobrerepresentación en la *muestra.* Por esta razón deberá eliminarse su duplicidad.

d) El *marco muestral* tampoco ha de incluir *unidades* que no correspondan a la *población* que se analiza. La inclusión de estas *unidades* reduce la probabilidad de elección de las *unidades* que sí pertenecen a la *población.*

EJEMPLO DE SUPRESIÓN EN UN MARCO DE MUESTREO

Si se hiciera una encuesta a la población mayor de 40 años, habría que circunscribir la *población* a esta cuota de edad. Las personas de 40 años y menos deberían eliminarse del *marco muestral.* Su supresión podría hacerse *a priori* (antes de proceder a la extracción de la *muestra), o a posteriori* (una vez que la *muestra* ha sido seleccionada). Esta segunda opción suele ser más practicada. Consiste en sustraer (de la *muestra* obtenida) aquellas *unidades* que no pertenezcan a la *población* de interés.

e) El *marco muestral* ha de contener información suplementaria que ayude a la localización de las *unidades* seleccionadas: teléfono, dirección.
f) Ante todo, el *marco* debería ser fácil de utilizar. Aquellos que están en soporte magnético simplifican y automatizan bastante el proceso de extracción muestral. No sólo reducen los costes del diseño de la *muestra* (en tiempo y en dinero), sino que también contribuyen a la reducción de errores (o *sesgos*) en la *selección de la muestra.*

Por todos los aspectos señalados, es importante que el investigador evalúe la calidad del *marco muestral* antes de comenzar la extracción de la *muestra.* El conocimiento de cómo se ha procedido en su elaboración y la periodicidad de su actualización, reducirá los *sesgos* motivados por su desconocimiento.

Por último, insistir en que la "representatividad" de la *muestra* sólo podrá ser de la *población* contenida en el *marco muestral* empleado. De ahí que se recomiende:

> "Cuando se informen resultados, el investigador debe decir a los lectores a quién se dio y a quién no se dio una oportunidad de ser seleccionado, cómo aquellos omitidos diferían de los otros y si había o no algunas personas muestreadas para quienes las oportunidades de selección no eran conocidas" (Fowler, 1988: 22).

5.2. El tamaño de la muestra

El número de unidades a incluir en la *muestra* constituye una de las decisiones preliminares en cualquier *diseño muestral.* En esta decisión participan diferentes factores. Éstos pueden resumirse en los seis siguientes:

a) El *tiempo* y los *recursos* disponibles para llevar a cabo la investigación.
b) La *modalidad de muestreo* seleccionada en orden a alcanzar los objetivos esenciales de la investigación.
c) La diversidad de los *análisis de datos* prevista.
d) La *varianza* o heterogeneidad poblacional.
e) El *margen de error* máximo admisible para la estimación de los *parámetros poblacionales*.
f) El *nivel de confianza* de la *estimación muestral*.

• *El tiempo y los recursos disponibles*

Sin duda, dos componentes esenciales en cualquier investigación son el *tiempo* y los *recursos* (económicos, materiales y humanos) concedidos al equipo investigador para la materialización del estudio propuesto.

En función de cuánta sea la dotación económica y los plazos temporales otorgados para cada fase de la investigación, el tamaño final de la *muestra* se verá aumentado o mermado.

• *La modalidad de muestreo seleccionada*

La elección de una (o varias) de las modalidades de *muestreo (probabilístico* o *no probabilístico)* se halla determinada –como ya se expuso en el Capítulo 3– por la confluencia de tres componentes claves en la práctica de cualquier investigación: los *objetivos*, los *recursos* y el *tiempo* dado para su realización. La incidencia de estos tres elementos no sólo influye en la selección de una *estrategia* concreta de investigación (y las consiguientes *técnicas de obtención y de análisis de datos),* sino también en la configuración del *diseño de la muestra.*

Como posteriormente se verá, las exigencias de *tamaño de la muestra* varían en consonancia con la modalidad elegida para la extracción de la *muestra.*

En general, los *diseños muestrales no probabilísticos* demandan un *tamaño muestral* inferior a los *diseños probabilísticos.* Estos últimos están más orientados al logro de *muestras* "representativas". Lo que exige aumentar el tamaño de la muestra para reducir el *error muestral.*

En ambas agrupaciones de diseños *(probabilísticos* y *no probabilísticos)* se observan también variaciones entre las diferentes modalidades de muestreo. Así, por ejemplo, el *tamaño muestral* que se requiere para la realización de un *muestreo aleatorio sistemático* es inferior al necesario para efectuar un *muestreo aleatorio estratificado* (que garantice la presencia en la *muestra* de determinados subgrupos de población). Lo mismo acontece si se decide aplicar un *muestreo por cuotas,* con preferencia a otros *muestreos no probabilísticos* (como el *estratégico* o el de *bola de nieve).*

- *La diversidad de los análisis de datos prevista*

Cuando se proyecta cualquier investigación, antes de *diseñar la muestra* y la *estrategia* a seguir en la obtención de información, hay que anticipar la variedad de *análisis* que se estimen necesarios para la consecución de los *objetivos* de la investigación.

Si el equipo investigador cree de interés la aplicación de alguna (o varias) *técnicas estadísticas multivariables,* deberá procurar que la *muestra* analizada incluya un número elevado de casos. Ello contribuye a la reducción de los errores de predicción y al aumento, consiguiente, del poder predictivo del modelo obtenido.

Para la realización de *análisis multivariables* se precisa una cierta proporcionalidad entre el número de observaciones *(tamaño de la muestra)* y el número de variables incluidas en el estudio. En *regresión múltiple,* por ejemplo, la proporción es, al menos, de 20 veces más casos que *variables independientes* (si se pretende que el modelo resultante sea *estadísticamente significativo).* Esta proporción se duplica (40 veces más casos que *variables independientes),* cuando se realiza un análisis de regresión "paso a paso" (Tabachnick y Fidell, 1989). En caso contrario, habrá que eliminar *variables independientes* o proceder a su agrupación. Para *análisis bivariados* (de cruces de variables) también se precisa un *tamaño muestral* elevado. Especialmente, si las *variables* comprenden un número amplio de *categorías.*

- *La varianza o heterogeneidad poblacional*

Si la *muestra* intenta ser una representación a pequeña escala de una *población* (o *universo),* su tamaño se verá afectado por la *varianza poblacional.* Cuanto más heterogénea sea la *población,* mayor será su *varianza poblacional.* En consecuencia, necesitará un mayor *tamaño muestral* para que la variedad de sus componentes se halle representada en la *muestra.*

Las exigencias de *tamaño muestral* se reducen, en cambio, en universos homogéneos, caracterizados por una baja *varianza poblacional.*

El conocimiento de la homogeneidad o heterogeneidad poblacional resulta, por tanto, primordial en la decisión del *tamaño de la muestra.* Rodríguez Osuna (1991) sintetiza en dos las formas esenciales de acceder a dicho conocimiento en la práctica de la investigación empírica:

a) La *experiencia* adquirida en estudios que se repiten con periodicidad (como los *Eurobarómetros* realizados mensualmente por la Comunidad Económica Europea (CEE); o la *Encuesta de Población Activa* del INE). Estos estudios proporcionan los valores de las *varianzas* para distintas variables. Lo que es de ayuda para el diseño muestral de futuras indagaciones empíricas.

Sin embargo, hay que matizar que ello es posible siempre y cuando ambas *poblaciones* (la del estudio previo y la de la investigación que se proyecta) coincidan.

b) La realización de *estudios pilotos* previos a la investigación principal, que ayuden al cálculo de las *varianzas* de las variables de interés.

Cuando se desconoce el valor de la *varianza poblacional* (situación muy habitual en la práctica de la investigación social), se recurre al supuesto más desfavorable: se toma el producto de las probabilidades "P" (de aparición de un suceso) y "Q" (que indica la no ocurrencia del suceso o evento; siendo su valor igual a "1 – *P*") como equivalente a la *varianza poblacional;* ambas probabilidades presentarían el valor de 0,50.

Este es el supuesto más desfavorable porque conlleva un *tamaño muestral* superior (con el consiguiente incremento de los costes de la investigación) al correspondiente cuando se conocen los valores reales de dichas proporciones en la población, y éstos se alejan del valor 0,50.

A modo de ejemplo, véase el Cuadro 5.1. En él se resumen algunos *tamaños muestrales* correspondientes a *poblaciones infinitas* (mayores de 100.000 unidades), al *nivel de confianza* más empleado (2 sigma = 95,5% de probabilidad de acertar en la estimación del *parámetro* poblacional), para distintos valores de "P" y "Q" (expresados en porcentajes), y *límites de error* establecidos.

CUADRO 5.1. Tamaño muestral para poblaciones infinitas a un nivel de confianza del 95,5% (2 sigma).

Límites de error (%) para $\pm 2\sigma$	*Valores presupuestos de P y Q (%)*					
	1/99	10/90	20/80	30/70	40/60	50/50
0,1	39.600	360.000	640.000	840.000	960.000	1.000.000
0,5	1.584	14.400	25.600	33.600	38.400	40.000
1,0	396	3.600	6.400	8.400	9.600	10.000
1,5	176	1.600	2.844	3.733	4.267	4.444
2,0	99	900	1.600	2.100	2.400	2.500
2,5	63	576	1.024	1.344	1.536	1.600
3,0	44	400	711	933	1.067	1.111
3,5	32	294	522	686	784	816
4,0	25	225	400	525	600	625
5,0	16	144	256	336	384	400

En el cuadro puede observarse que el tamaño muestral aumenta, conforme los valores de "P" y "Q" se aproximan al valor 0,50 (indistintamente de cuál sea el *margen de error* fijado). Así, por *ejemplo*, para un *margen de error* de ±2%, el *tamaño muestral* para una *población infinita* será de 900 unidades, si $P = 10$ y $Q = 90$. En cambio, el *tamaño de la muestra* será más del doble (2.500 unidades), cuando $P = 50$ y $Q = 50$.

Este incremento considerable en el *tamaño de la muestra* reduce, no obstante, las posibilidades de equivocación en la estimación de los *parámetros poblacionales*

correspondientes a las variables del estudio. Razón por la cual se adopta este supuesto, siempre que se desconozca el valor de la *varianza poblacional* respecto a las variables que se analizan.

Los *tamaños muestrales* que figuran en el Cuadro 5.1 se han obtenido a partir de la fórmula común para el cálculo del *tamaño muestral* para *universos infinitos,* a un *nivel de confianza* de 2 sigma (que supone un 95,5% de probabilidad de acertar en la estimación):

$$N = \frac{4\hat{P}\hat{Q}}{E^2}$$

donde "E" representa el *error muestral.*

Las fórmulas del *tamaño muestral* se comentan al final de este apartado.

• *El margen de error máximo admisible*

En el Cuadro 5.1 también puede observarse la correspondencia entre las variables *error* y *tamaño muestral:* los incrementos en el *tamaño de la muestra* repercuten en una mayor precisión en la *estimación* de los *parámetros poblacionales,* con la consiguiente reducción del *error muestral.* En cambio, en *muestras* pequeñas, el *error de muestreo* aumenta, manteniendo constante la *varianza poblacional.*

La Figura 5.1 representa, de forma gráfica, la relación negativa existente entre las variables *error* y *tamaño muestral.* Para su elaboración se han tomado los valores que aparecen en el Cuadro 5.1 correspondientes a la situación habitual de desconocimiento del valor de la *varianza poblacional* ($P = Q = 0{,}50$).

Como puede observarse en la Figura 5.1, a medida que aumenta el volumen del *tamaño de la muestra,* se produce un decrecimiento en el valor del *error muestral.* Si el *tamaño de la muestra* se amplía, por *ejemplo,* de 816 a 2.500 unidades, el *error muestral* disminuye en un 1,5% (pasando de representar un 3,5% a sólo un 2%).

También se advierte que a partir del 2% de *error,* se disparan los crecimientos en el *tamaño de la muestra* para alcanzar una mínima ganancia en la reducción del *error muestral.* Pasar del 2% al 1% de *error* supone un aumento de 2.500 a 10.000 unidades, con el incremento correspondiente en los costes (económicos y temporales) de la investigación; a cambio, sólo se logra una leve mejora en la precisión de la *estimación del parámetro poblacional* (un 1%). Por lo que, el investigador deberá encontrar un punto intermedio entre el *tamaño* y el *error muestral,* ante la tendencia observada en muestras grandes a proporcionar mínimos incrementos en adecuación en la *estimación de parámetros.*

El *error muestral* interviene en la decisión sobre el *tamaño de la muestra* siempre que el *diseño muestral* sea *probabilístico.* En este caso, el investigador fija el *error a priori,* sopesando la precisión que desea para sus *estimaciones,* con los costes que supondría la reducción del *error muestral.* Los *errores* comprendidos entre el 2,5% y el 2% son los más frecuentes en la investigación social.

Figura 5.1. Representación gráfica de la relación entre las variables error y tamaño muestral para $P = Q = .50$ y nivel de confianza del 95,5%.

- *El nivel de confianza de la estimación*

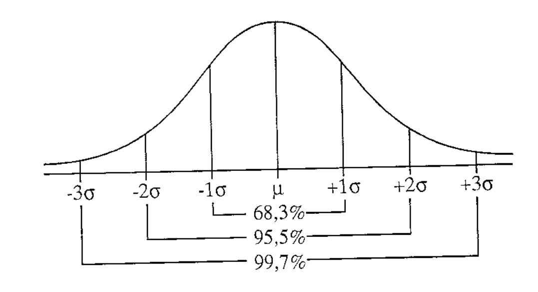
Un último elemento a considerar en la decisión sobre el *tamaño de la muestra* –aunque también sólo en los *diseños probabilísticos*– es el *nivel de confianza* que el investigador concede a sus *estimaciones*.

El *nivel de confianza* expresa –como su nombre indica– el grado de "confianza" (o de "probabilidad") que el investigador tiene en que su estimación se ajuste a la realidad.

Tres son los *niveles de confianza* comunes en la investigación social. Corresponden a áreas bajo la curva normal acotadas por distintos valores de *desviación típica* (denominada *sigma* (σ) en referencia a la *desviación poblacional*). De ellos el más habitual es 2σ, que supone un 95,5% de probabilidad de acertar en la estimación a partir de los datos muestrales.

La *distribución normal* se aplica en la *estadística inferencial* para la estimación de la probabilidad de que un determinado evento acontezca. Representa una curva perfectamente simétrica, en forma de campana, y que admite infinitos valores (unidades "Z": *unidades de desviación típica*). El área total bajo la *curva normal* es 1 (dado que la *probabilidad* siempre es un valor comprendido entre 0 y 1). En función de cuál sea el valor de "Z", variará la *probabilidad* concedida al evento en cuestión (véase la tabla del área bajo la curva normal en el Anexo).

Valores de la distribución normal		*Probabilidad comprendida en el intervalo*
Entre	±1σ	0,683
	±2σ	0,955
	±3σ	0,997



Si se toma una unidad de *desviación típica*, la *probabilidad* de acertar en la estimación sería del 68,3%. Si en vez de una se toman dos unidades de *desviación*, la *probabilidad* de acierto se incrementa hasta el 95,5% (habiendo sólo un 4,5% de *probabilidad* de equivocación en la *estimación del parámetro poblacional).*

La plena seguridad en la estimación (exactamente el 99,7%) se adquiere, si se fijan en tres las unidades de *desviación típica* (3 sigma). Pero, no es éste el *nivel de confianza* usual en la práctica investigadora, sino el 95,5% (equivalente a 2). Ante ello, la pregunta inmediata es ¿por qué?

Para encontrar la respuesta, véase el Cuadro 5.2. En él se ejemplifica cómo (para una misma *varianza poblacional)* varía el *tamaño muestral,* dependiendo de cuál sea el *nivel de confianza* adoptado. Incrementos mínimos en *nivel de confianza* (del 95,5% de probabilidad de acierto al 99,7%) supone añadir más del doble de unidades a la *muestra.*

Si para un *error* del 2% y *varianza* P = Q = 50, el *tamaño muestral* sería de 2.500 unidades (para un *nivel de confianza* de 95,5%), al aumentar el nivel de confianza al

CUADRO 5.2. Tamaño muestral para poblaciones infinitas a un nivel de confianza del 95,5% y del 99,7%.

Límites de error (%)	*Nivel de confianza (%)*	*Valores presupuestos P Y Q (%)*				
		10/90	20/80	30/70	40/60	50/50
1,0	95,5 99,7	3.600 8.100	6.400 14.400	8.400 18.900	9.600 21.600	10.000 22.500
2,0	95,5 99,7	900 2.025	1.600 3.600	2.100 4.725	2.400 5.400	2.500 5.627
2,5	95,5 99,7	576 1.296	1.024 2.304	1.344 3.024	1.536 3.456	1.600 3.600
3,0	95,5 99,7	400 900	711 1.600	933 2.100	1.067 2.400	1.111 2.500
4,0	95,5 99,7	225 506	400 900	525 1.181	600 1.350	625 [illegible]

99,7%, la muestra pasaría a estar integrada por 5.627 elementos. La ganancia en precisión es ínfima, mientras que los costes de la investigación (en tiempo y dinero) supondrían más del doble de los presupuestados para un nivel de confianza del 95,5%.

Todos los aspectos referidos participan en el cálculo del *tamaño de una muestra probabilística*. La *fórmula* genérica para una *muestra aleatoria (simple o sistemática)* sería la siguiente, cuando el universo o población estuviese compuesto por más de 100.000 unidades:

$$n = \frac{Z^2 \hat{S}^2}{E^2} \quad \text{o} \quad n = \frac{Z^2 \hat{P}^2 (1 - \hat{P})}{E^2}$$

donde: "Z" representa las *unidades de desviación típica* correspondientes al *nivel de confianza* elegido (2σ o 3σ, fundamentalmente).

"$\hat{S}^2$ constituye el valor de la *varianza poblacional*. Este equivale al producto de las proporciones $\hat{P}$ y $\hat{Q}$ siendo $\hat{Q} = 1 - \hat{P}$.

"E" denota el *error* máximo permitido que el investigador establece *a priori*.

Si el universo estuviese integrado por 100.000 unidades o menos, se trataría de una *población finita*. En este caso, habría que introducir un factor de corrección, quedando las fórmulas transformadas de la siguiente manera:

$$n = \frac{Z^2 \hat{S}^2 N}{E^2 (N - 1) + Z^2 \hat{S}^2} \quad \text{o} \quad n = \frac{Z^2 \hat{P} \hat{Q} N}{E^2 (N - 1) + Z^2 \hat{P} \hat{Q}}$$

donde: "N" es el tamaño de la población.

En *muestras aleatorias estratificadas* y *por conglomerados* se introducen otras variaciones que se comentarán cuando se describan ambas modalidades de muestreo (apartado 5.4).

EJEMPLOS DE CÁLCULO DEL TAMAÑO MUESTRAL

Se desea conocer el tamaño de la muestra necesario para una encuesta a la población mayor de 18 años, con objeto de medir el voto en las próximas elecciones municipales. El error máximo permitido será ±2,5%.

a) Si la encuesta se efectúa en Madrid capital, sin información previa sobre el porcentaje de voto. Para un nivel de confianza de 95,5% (2σ), el tamaño muestral necesario sería:

$$n = \frac{Z^2\hat{P}Q}{E^2} = \frac{2^2 \cdot 50 \cdot 50}{2{,}5^2} = 1.600 \text{ unidades}$$

En cambio, para un nivel de confianza del 99,7% (3σ), el tamaño muestral sería más del doble:

$$n = \frac{Z^2\hat{P}\hat{Q}}{E^2} = \frac{3^2 \cdot 50 \cdot 50}{2{,}5^2} = 3.600 \text{ unidades}$$

b) Si se parte de la predicción (a partir de sondeos realizados con anterioridad) de que el PSOE puede alcanzar el 24% del total de votos:

$$n = \frac{Z^2\hat{P}\hat{Q}}{E^2} = \frac{2^2 \cdot 24 \cdot 76}{2{,}5^2} = 1.167 \text{ unidades para } 2\sigma$$

$$n = \frac{Z^2\hat{P}\hat{Q}}{E^2} = \frac{3^2 \cdot 24 \cdot 76}{2{,}5^2} = 2.627 \text{ unidades para } 3\sigma$$

c) Si la encuesta se llevase a cabo en el municipio de Torrelodones, con una población de hecho (censada en 1991) de 7.113 individuos.

$$n = \frac{Z^2\hat{P}\hat{Q}N}{E^2(N-1) + Z^2\hat{P}\hat{Q}} = \frac{4 \cdot 50 \cdot 50 \cdot 7.113}{(2{,}5^2 \cdot 7.112) + (4 \cdot 50 \cdot 50)} = 1.306 \text{ para } 2\sigma$$

$$n = \frac{Z^2PQN}{E^2(N-1) + Z^2\hat{P}\hat{Q}} = \frac{9 \cdot 50 \cdot 50 \cdot 7.113}{(2{,}5^2 \cdot 7.112) + (9 \cdot 50 \cdot 50)} = 2.390 \text{ para } 3\sigma$$

5.3. El error muestral

Cuando se diseña una *muestra*, el objetivo primordial es conseguir un elevado nivel de adecuación en la selección de la *muestra*, respecto de la *población* a la que pertenece. Ello contribuye a que la investigación adquiera *validez externa* (como ya se vio en el Capítulo 3).

Pero, por muy perfecta que sea la *muestra*, como únicamente se analiza una parte de la *población* (y esa *muestra* sólo representa una de todas las posibles *muestras* que pueden extraerse de una misma población), siempre habrá alguna divergencia entre los valores obtenidos de la *muestra (estimaciones)* y los valores correspondientes en la *población (parámetros)*. Esa disparidad se denomina *error muestral*: el grado de inadecuación existente entre las *estimaciones muestrales* y los *parámetros poblacionales*.

Como todos los resultados muestrales se hallan sujetos a *errores de muestreo,* cuando se analizan datos muestrales, debería comprobarse la *representatividad* de la *mues-*

tra e identificar (si los hubiese) los *sesgos* existentes. Ello proporciona un mayor rigor a las conclusiones que se deduzcan de la investigación efectuada.

Si la extracción de la *muestra* se ha llevado a cabo de forma aleatoria *(muestreo probabilístico),* podrá calcularse el *error muestral,* con la ayuda de la *teoría de la probabilidad;* en caso contrario *(muestreo no probabilístico),* no. Esta constituye una de las ventajas importantes de los *diseños muestrales probabilísticos:* la capacidad de determinar la precisión de las *estimaciones muestrales.*

Para el cálculo del *error muestral* (en *muestras probabilísticas,* exclusivamente), se acude al estadístico llamado *"error típico".* En la *teoría de la probabilidad,* este estadístico mide la extensión a la que las *estimaciones muestrales* se distribuyen alrededor del *parámetro poblacional.* Concretamente, se especifica que aproximadamente el 68% de las *estimaciones muestrales* se hallarán comprendidas entre ±1 vez el *error típico del parámetro poblacional;* el 95,5%, entre ±2 veces el *error típico;* y, finalmente, el 99,7%, entre ±3 veces el *error típico.*

En suma, el *nivel de confianza* en la *estimación* aumenta conforme se amplia el *margen de error.* El *nivel de confianza* más utilizado es -como ya se indicó en el apartado anterior- 2σ (sigma). Este nivel de confianza expresa que hay un 95,5% de probabilidad de que la *estimación muestral* se halle comprendida en el intervalo definido por dos veces el *error típico del parámetro poblacional.*

En el cálculo del *error típico* intervienen los elementos siguientes:

a) El *tamaño muestral.*
b) La *varianza poblacional.*
c) El *nivel de confianza* adoptado.
d) El *tipo de muestreo* realizado.

- *El tamaño muestral*

No es la proporción de la población que constituye la muestra (la *fracción de muestreo* = n/N) lo que determina el *error muestral* (salvo si la *fracción* resulta elevada), sino el *tamaño de la muestra.* A medida que aumenta el *tamaño de la muestra,* decrece el *error muestral (cualquiera que sea el valor de la* varianza poblacional).

Dicha afirmación puede comprobarse en el Cuadro 5.3. En él figuran los *errores típicos* correspondientes a determinados *tamaños muestrales.* Estos *errores típicos* disminuyen conforme se incrementa el *tamaño muestral,* indistintamente de los valores presupuestos de P y Q. Incluso, véase que para reducir el *error típico* a la mitad, habría que cuadruplicar el *tamaño de la muestra.*

Por *ejemplo,* una *muestra* compuesta de 500 unidades en la situación más desfavorable (P = Q = 50) presenta un *error típico* de ±4,5% (a un n. c. del 95,5%). Para reducir dicho *error* a ±2,2%, habría que multiplicar el *tamaño muestral* por 4, obteniéndose una muestra de 2.000 unidades. Si el *tamaño muestral* se amplía de 1.000 a 4.000 unidades, el *error* también se reduce a la mitad (de ±3,2% a ±1,6%).

CUADRO 5.3. Errores típicos correspondientes a distintos tamaños muestrales a un nivel de confianza del 95,5% (2 σ).

Número de elementos muestra	*Valores presupuestos de P y Q (%)*					
	1/99	10/90	20/80	30/70	40/60	50/50
50	2,8	8,5	11,4	13,0	13,9	14,2
100	2,0	6,0	8,0	9,2	9,8	10,0
200	1,4	4,3	5,7	6,5	7,0	7,1
500	0,9	2,7	3,6	4,1	4,4	4,5
1.000	0,6	1,9	2,6	2,9	3,1	3,2
1.500	0,5	1,6	2,1	2,4	2,5	2,6
2.000	0,4	1,3	1,8	2,0	2,2	2,2
3.000	0,4	1,1	1,5	1,7	1,8	1,8
4.000	0,3	1,0	1,3	1,4	1,5	1,6
5.000	0,3	0,9	1,1	1,3	1,4	1,4
10.000	0,2	0,6	1,8	1,0	1,0	1,0
50.000	0,1	0,2	0,3	0,4	0,4	0,4

- *La varianza poblacional*

El nivel de *heterogeneidad* de una *población* (respecto a una variable considerada), favorece el *error muestral,* excepto si se aumenta el *tamaño muestral* para incluir todas las distintas variedades que componen el *universo.*

El *error muestral* se halla más presente en *poblaciones* heterogéneas (de elevada *varianza poblacional)* que en *universos homogéneos.*

- *El nivel de confianza adoptado*

Como aumentar el *nivel de confianza* supone ampliar el *tamaño de la muestra,* obviamente un mayor *nivel de confianza* llevará consigo la reducción del *error muestral.* Pese a ello, hay que insistir en que, cuando se diseña una *muestra,* se valoran conjuntamente distintos aspectos, buscando el equilibrio entre precisión y coste.

Incrementos en el *tamaño de la muestra* conllevan una ampliación del *nivel de confianza* en la *estimación muestral* (descendiendo el *error).* Pero esta mejora en precisión repercute en unos costes superiores de la investigación. De ahí que el *nivel de confianza* más empleado sea el ya referido de 2σ (95,5% de *probabilidad* de acierto en la *estimación*).

- *El tipo de muestreo realizado*

El *error muestral* también se halla afectado por el procedimiento de selección de las unidades muestrales. En general, el *muestreo aleatorio estratificado* es (de los di-

seños muestrales probabilísticos) el que genera un menor *error muestral.* En cambio, el *muestreo aleatorio por conglomerados* es (de los *probabilísticos)* el que ocasiona un mayor *error muestral.* Aunque la agrupación de la *muestra* en *conglomerados* presenta la gran ventaja de reducir los costes (en tiempo y dinero) del trabajo de campo, éste a su vez repercute en una desventaja importante: incrementa el *error típico de la muestra,* con la consiguiente disminución en precisión de los estimaciones de la población.

A estas afirmaciones genéricas cabe añadir –de acuerdo con Fowler (1988)– las precisiones siguientes:

a) El *muestreo aleatorio sistemático* ocasiona *errores de muestreo* equivalentes a las *muestras aleatorias simples,* siempre y cuando no exista ninguna estratificación en la muestra.
b) Las *muestras aleatorias estratificadas* presentan menores *errores de muestreo* que las *muestras aleatorias simples* del mismo tamaño; pero, sólo si existe heterogeneidad entre los estratos.
c) El *muestreo aleatorio por conglomerados* genera un mayor *error típico,* sobre todo cuanto mayor es el tamaño de los conglomerados en la última fase de muestreo.

Dadas estas variaciones en el *error muestral* debidas al diseño realizado, los cálculos del *error típico* han de adecuarse a las fórmulas apropiadas a cada *diseño muestral.*

Para una *muestra aleatoria simple* o *sistemática,* las fórmulas correspondientes al error típico (E) serían las siguientes:

	Universo infinito	*Universo finito (≤100.000 unidades)*
Error típico de la media	$E_{(\bar{x})} = \sqrt{\frac{\hat{S}^2}{n}}$	$E_{(\bar{x})} = \sqrt{\frac{\hat{S}^2}{n}\frac{N-n}{N-1}}$
Error típico de una proporción	$E_{(p)}\sqrt{\frac{\hat{P}\hat{Q}}{n}}$	$E_{(p)}\sqrt{\frac{\hat{P}\hat{Q}}{n}\frac{N-n}{N-1}}$

Las proporciones "P" y "Q" pueden expresarse tanto en porcentajes como en tantos por uno, obteniéndose los mismos resultados.

En todos los casos, los resultados obtenidos se multiplicarían por el *nivel de confianza* adoptado. Si es 95,5%, por 2σ; y si fuera 99,7%, por 3σ.

A partir de los valores obtenidos, se calculan los *intervalos de confianza* para el *nivel de probabilidad* fijado. Ello posibilita conocer cuánto se aproxima la estimación muestral al *parámetro poblacional.*

EJEMPLOS DE ERROR MUESTRAL

a) Un equipo de investigadores desea conocer los hábitos de consumo de la población juvenil española. Para ello entrevistan (mediante *cuestionario)* a 2.000 jóvenes distribuidos por toda España. Del estudio se concluye (entre otros aspectos) que la media de gastos mensuales es de 30.000 pts al mes, con una desviación típica de 5.600. Calcular el error de la estimación muestral. A partir de él, inferir cuál será el parámetro poblacional correspondiente a un nivel de probabilidad de 2σ.

$$E_{(\bar{x})} = \sqrt{\frac{\hat{S}^2}{n}} = \sqrt{\frac{5.600^2}{2.000}} = 125,21$$

Si el *error* se multiplica por 2σ, se realiza una estimación del *intervalo de confianza* que comprende el *parámetro poblacional,* sumando y restando a la estimación muestral el producto del *error* por el *nivel de confianza* elegido: 30.000 ± (2)(125,21).

En consecuencia, hay un 95,5% de probabilidad de que la media de gastos mensuales de la población juvenil española se halle comprendida entre 29.750 y 30.250 pesetas al mes.

Si el *nivel de confianza* elegido hubiera sido 3σ, el *intervalo de confianza* sería ligeramente más amplio; por lo que sería menor la probabilidad de equivocación en la estimación. Los *límites de confianza* serían, respectivamente: 30.000 – (3)(125,2) y 30.000 + (3)(125,2). Es decir, habría un 99,7% de probabilidad de que la media de gastos mensuales de la población juvenil española se halle comprendida entre 29.624 y 30.376 pesetas al mes.

b) De los 2.950 menores internados en un centro de reforma se ha extraído una muestra de 780 menores de distintas edades. De ellos, el 49% fueron acusados de hurto. ¿Cuál será la proporción de detenidos por hurto en la población total recluida en dicho centro a un nivel de confianza de 2σ?

$$E_{(p)} = \sqrt{\frac{\hat{P}\hat{Q}}{n}\,\frac{N-n}{N-1}} = \sqrt{\frac{(49)(51)}{780}\,\frac{2950-780}{2950-1}} = 1,54\%$$

Los *límites de confianza* serían, respectivamente: 49 – (2)(1,5) y 49 + (2)(1,5). Por tanto habría un 95,5% de probabilidad de que el procentaje de menores detenidos por hurto (en todo el centro) estuviese comprendido entre el 46 y el 52% de la población total.

Si la muestra fuese *aleatoria estratificada proporcional,* se introducirían las modificaciones siguientes en las fórmulas del *error típico:*

Error típico de la media: $$E_{(\bar{x}prop)} = \sqrt{\frac{\sum_{i=1}^{n} n_i \hat{S}_i^2}{n^2}}$$

Error típico de una proporción:

$$E_{(Pprop)} = \sqrt{\frac{\sum_{i=1}^{n} n_i \hat{P}_i \hat{Q}_i}{n^2}}$$

donde "$\hat{P}_i$" es la proporción de la muestra en el estrato i que posee el atributo en cuestión.
"$\hat{Q}$"= $1 - \hat{P}_i$": la proporción que no posee dicho atributo.
"$\hat{S}_i^2$" = $\hat{P}_i * \hat{Q}_i$": estimación de la *varianza* de la variable de interés para la población en el estrato i.
"Σ": sumatorio de todos los estratos, desde 1 hasta n.
"n_i": tamaño de la muestra total.

EJEMPLO DE ESTIMACIÓN DEL ERROR MUESTRAL EN UNA MUESTRA ALEATORIA ESTRATIFICADA

Se quiere conocer el grado de aceptación de los nuevos planes de estudio por la población universitaria española. A tal fin se escoge una *muestra* integrada por 2.439 alumnos de la Universidad Complutense. La *muestra* se *afija* (distribuye) entre las distintas facultades de manera *proporcional* al número de alumnos registrados en dichas facultades. Para ello se multiplica la proporción que representa cada uno de los *estratos* (facultades diferenciadas) –columna (2)– por el *tamaño muestral* (2.439), obteniéndose las cifras que aparecen en la columna (3). En la columna (4) figuran los porcentajes de alumnos que en cada una de las facultades afirman estar satisfechos con los nuevos planes de estudio.

Facultades	*(1) Población*	*(2) Pobl. prop.*	*(3)* n_i	*(4)* P_i	*(5)* $n_iP_i(1-P_i)$	*(6)* $r_i=n_iP_i$
Medicina	4.535	0,046	112	0,49	27,99	55
Farmacia	4.983	0,050	122	0,35	27,76	43
Veterinaria	2.682	0,027	66	0,49	16,49	32
Derecho	18.600	0,188	459	0,71	94,51	326
Económicas/Empresa.	13.755	0,139	339	0,83	47,83	281
Políticas/Sociología	6.888	0,070	171	0,42	41,66	72
Psicología	4.567	0,046	113	0,43	27,70	49
Filol./Filosofía/Pedag.	10.613	0,107	261	0,39	62,09	102
Geografía/Historia	6.063	0,061	149	0,51	37,24	76
CC. Información	13.900	0,141	344	0,59	83,21	203
Química/Fisica	5.568	0,056	137	0,69	29,30	95
CC. Exactas	2.411	0,024	59	0,47	14,70	28
Biología/Geología	4.388	0,044	107	0,44	26,36	47
	98.953		2.439		536,84	1.409

Para estimar el *error típico* de esta *muestra aleatoria estratificada proporcional*, se procede a multiplicar el número de alumnos que componen la muestra en cada estrato (columna 3) por la proporción de alumnos a favor de los nuevos planes de estudio (columna 4) y por la proporción de alumnos en contra ($\hat{Q} = 1 - \hat{P}$). Como resultado se obtienen las cifras que aparecen en la columna (5). Después, se suman dichos productos y se sustituyen en la fórmula correspondiente:

$$E_{(pprop)} = \sqrt{\frac{\sum_{i=1}^{n} n_i \hat{P}_i \hat{Q}_i}{n^2}} = \sqrt{\frac{536{,}84}{2.439^2}} = 0{,}0095$$

Para demostrar que el *muestreo estratificado* genera un menor *error típico* que el *muestreo aleatorio simple*, se calcula el *error típico* que se tendría si la *muestra* hubiera sido *aleatoria simple*. Con este propósito se estima la proporción de los alumnos encuestados que afirman estar a favor de los nuevos planes de estudio. En la columna (6) figura el número de alumnos que contestaron estar a favor en cada uno de los estratos considerados. Si se suman dichas cantidades (r_i) y se divide por el total de sujetos analizados (2.439), se obtiene la proporción global de alumnos a favor: el 57,8% de los encuestados.

$$P = \frac{\Sigma r_i}{n} = \frac{1.409}{2.439} = 0{,}578$$

Por tanto

$$E_{(p)}\sqrt{\frac{\hat{P}\hat{Q}}{n}\frac{N-n}{N-1}} = \sqrt{\frac{(0{,}58)(0{,}42)}{2.439}\frac{(98.953-2.439)}{(98.953-1)}} = 0{,}0099$$

En este caso, la diferencia entre ambos tipos de errores parece mínima. Pero, si se eleva el *error típico* al cuadrado *(varianza)* obtenido del *muestreo aleatorio estratificado*, y se divide por el cuadrado del *error típico* que se obtendría mediante un *muestreo aleatorio simple*, se comprueba el *efecto del diseño* (deff), enunciado por Kish (1965):

$$deff = \frac{(0{,}0095)^2}{(0.0099)^2} = 0{,}921$$

De esta forma se concluye que la *varianza* de la *muestra estratificada proporcionalmente* es un 8% más pequeña que la *varianza* de la *muestra aleatoria simple*.

Por tanto, si mediante un *muestreo aleatorio simple*, se quisiera alcanzar la misma precisión en la estimación de los parámetros poblacionales, habría que ampliar el *tamaño muestral* a 2.648 (esta cifra se obtiene dividiendo 2.439/0,921). Con ello se demuestra la mayor precisión alcanzada mediante un *muestreo aleatorio estratificado*. La diferencia entre ambos tipos de errores crece a medida que aumentan las variaciones observadas en los estratos respecto a la variable considerada.

Por último, si la *muestra* fuese por *conglomerados*, la fórmula correspondiente al *error típico* sería la siguiente:

$$E_{(\bar{x}congl)} = \sqrt{\left(1 - \frac{m}{M}\right)\frac{Sb^2}{m}}$$

donde: *"M"*: número de *conglomerados* en la *población*.
"m": número de *conglomerados* seleccionados en la *muestra*.
"Sb²": la *varianza* de los valores del conglomerado x_i. Se obtiene:

$$Sb^2 = \frac{1}{m-1}\sum_{i=1}^{n}\left(\bar{x}_i - \bar{x}\right)^2$$

EJEMPLO DE ESTIMACIÓN DEL ERROR EN UNA MUESTRA ALEATORIA POR CONGLOMERADO

Adaptemos el ejemplo anterior a una *muestra por conglomerados*. Con este propósito se divide la población universitaria por curso académico y grupo. Imaginemos que fuesen 780 los grupos existentes en total. De ellos se decide escoger aleatoriamente 20 grupos *(conglomerados)* de 122 alumnos cada uno. De esta forma se obtendría una *muestra* total de 2.440 alumnos.

El *error típico* para esta *muestra por conglomerados* sería:

$$E_{(\bar{x}congl)} = \sqrt{\left(1 - \frac{m}{M}\right)\frac{Sb^2}{m}}$$

donde: *"M"*: número de *conglomerados* en la *población* (780).
"m": número de *conglomerados* seleccionados en la muestra (20).
"Sb²": la *varianza* de los valores del *conglomerado* a favor de los nuevos planes de estudio (0,099).

Entonces $$E_{(\bar{x}congl)} = \sqrt{\left(1 - \frac{20}{780}\right)\frac{0.099}{20}} = 0{,}0698$$

Por tanto, se habría obtenido un *error típico* superior al alcanzado mediante la *estratificación* (0,0095) y el *muestreo aleatorio simple* (0,0099), también.

En general, las estimaciones de los *errores típicos* para *muestras aleatorias por conglomerados* (especialmente si el *diseño muestral* es *polietápico)* resultan bastante complejas; si bien, en la actualidad, algunos paquetes informáticos –como el SAS *(Statistical Analysis System)*– calculan los *errores típicos* para cualquier variedad de *diseño muestral*.

5.4. Tipos de muestreo: diseños muestrales probabilísticos y no probabilísticos

La *representatividad* de una *muestra* no sólo depende de la magnitud de su *tamaño,* sino también del procedimiento seguido para la selección de las unidades muestrales. Las modalidades de muestreo son variadas, aunque cabe agruparlas en dos amplias categorías (muestreo *probabilístico* y *no probabilístico),* dependiendo de si el azar interviene en todas las fases de la selección.

• *Muestreo probabilístico o aleatorio*

Se fundamenta en la *aleatorización* como criterio esencial de selección muestral. Ello favorece que:

a) Cada unidad de la población tenga una *probabilidad* igual (y conocida *a priori)* de participar en la muestra.
b) La elección de cada unidad muestral sea independiente de las demás.
c) El cálculo de la adecuación de la estimación muestral *(error muestral)* a los parámetros poblacionales pueda hacerse dentro de unos márgenes de probabilidad específicos.

• *Muestreo no probabilístico*

A diferencia del anterior, la extracción de la *muestra* se efectúa siguiendo criterios diferentes de la *aleatorización* (como la conveniencia u otros criterios subjetivos). Ello da cabida a cualquier discreción por parte del equipo investigador. Además, repercute en:

a) La desigual *probabilidad* de las unidades de la población para formar parte de la muestra.
b) La dificultad de calcular el *error muestral.*
c) La introducción de *sesgos* en el proceso de elección muestral. Esto redunda en riesgos superiores de invalidez de los hallazgos de la investigación.

No obstante, el *muestreo no probabilístico* presenta dos *ventajas* notorias que le hacen atractivo en la práctica investigadora:

a) No precisa de la existencia de un *marco de muestreo.*
b) Su materialización resulta más sencilla y económica que los *muestreos probabilísticos.*

La elección entre un tipo de muestreo u otro vendrá condicionada por la confluencia de cuatro aspectos básicos:

a) La *dotación económica* de la investigación.
b) El período de *tiempo* programado para su ejecución.
c) La existencia de un *marco muestral* válido, que haga factible la *selección muestral aleatoria*.
d) El *grado de precisión* que el investigador quiera dar a la indagación.

El *muestreo probabilístico* se adecúa más a *propósitos* de:

a) Estimación de *parámetros*.
b) Comprobación de *hipótesis (test de significatividad)*.

En cambio, el *muestreo no probabilístico* se muestra más apropiado para:

a) La indagación *exploratoria (estudios piloto)*.
b) Estudios *cualitativos*, más interesadas en profundizar en la información aportada que en su representatividad estadística.
c) Investigaciones sobre población "marginal" (como prostitutas, delincuentes, drogadictos, homosexuales, inmigrantes ilegales, e incluso miembros menos "marginales" como parejas cohabitantes, por ejemplo), de difícil registro y localización. Lo que complica la aplicación de *diseños muestrales probabilísticos*.

El Cuadro 5.4 resume la variedad de *procedimientos muestrales* clasificados en "probabilísticos" y "no probabilísticos". Si bien, hay que precisar que éstos no forman compartimentos estancos. Un mismo *diseño muestral* puede compaginar diferentes variedades de muestreo en fases distintas de su desarrollo.

CUADRO 5.4. Modalidades de muestreo.

Probabilísticos	*No probabilísticos*
Simple Sistemático Estratificado Por conglomerados De rutas aleatorias	Por cuotas Estratégico "Bola de nieve"

5.4.1. Muestreo aleatorio simple

Constituye el prototipo de muestreo, en referencia al cual se estiman las fórmulas básicas para el cálculo del *tamaño* y del *error muestral.*

Como *muestreo probabilístico,* su realización exige la existencia de un *marco muestral* que cumpla las condiciones expuestas en el apartado 5.1. Una vez localizado, se asigna a cada unidad de la población (en el listado) un *número de identificación* (si éste no figura anexo al *marco muestral,* siguiendo un orden consecutivo) para, posteriormente, proceder a la extracción aleatoria de los integrantes de la *muestra.*

La selección muestral puede hacerse siguiendo una *tabla de números aleatorios,* mediante un programa de ordenador, u otro procedimiento que garantice que:

a) Cada unidad de la población tenga una probabilidad igual de participar en la muestra. Esta probabilidad viene determinada por la *fracción de muestreo* = n/N (el cociente entre el *tamaño de la muestra* y el total de la *población).*
b) La selección muestral sea totalmente *aleatoria* hasta alcanzar el *tamaño muestral* fijado. No obstante, se aconseja extraer un número mayor de unidades que las prefijadas en el *tamaño muestral* para las sustituciones (por imposibilidad de localizar a algunos de los integrantes originales de la *muestra,* o porque estos se nieguen a colaborar en la investigación, entre otras razones).

La elección de las unidades muestrales puede hacerse *sin reemplazamiento* (cuando cada unidad de la población puede ser elegida sólo una vez) o *con reemplazamiento* (si participa en elecciones sucesivas). Lo habitual es que el proceso de selección se efectúe *sin reemplazamiento.*

Cuando el *marco muestral* se halla en soporte magnético (como el último censo de población -de 1991-, por ejemplo), se facilita la selección muestral mediante ordenador. En este caso, el programa de ordenador configurado ejecuta todas las tareas correspondientes:

a) Numera a cada uno de los integrantes del *marco muestral* (o listado de las unidades de la población).
b) Genera su propia serie de *números aleatorios.*
c) Selecciona e imprime la *lista* de las unidades muestrales finalmente elegidas.

Si el uso del ordenador no resulta viable, se recurre al *procedimiento tradicional:* utilizar una *tabla de números aleatorios.* Estas *tablas* comprenden múltiples combinaciones de números extraídos al azar (o por sorteo), de forma que cada dígito presente una probabilidad igual de figurar en cualquier punto de la tabla. La actuación, en este caso, sería:

a) Elegir aleatoriamente un *punto de partida:* una columna o una fila cualquiera de la *tabla*. Puede comenzarse por cualquier lugar y moverse en cualquier dirección (hacia arriba, hacia abajo, a la derecha o a la izquierda). La única condición que se exige es que la elección no responda a un conocimiento previo de la distribución de los números en la *tabla*. Ello introduciría *sesgos* en la *muestra*.

b) El número de dígitos extraídos de la *tabla* ha de corresponder con el número de dígitos de la población contenida en el *marco muestral*.

Por *ejemplo*, si el *marco* estuviese compuesto por 10.000 unidades (5 dígitos), habría que numerar a cada unidad de la población con 5 dígitos: desde el primero (00001) hasta el último (10.000), correlativamente. A continuación, se escogerían números con 5 dígitos, desconsiderándose aquellas cifras que superen la última unidad poblacional numerada en el listado (10.000).

c) El individuo al que pertenezca el número extraído pasará a formar parte de la *muestra*, salvo que en el *marco* no se adjunte un medio para su localización (dirección postal, teléfono).

EJEMPLO DE MUESTREO ALEATORIO SIMPLE

Supóngase que va a realizarse una encuesta a licenciados en Ciencias Políticas y Sociología. El equipo investigador elige tomar como *marco de muestreo* el directorio de "colegiados" (en el Ilustre Colegio Nacional de Doctores y Licenciados en Ciencias Políticas y Sociología de Madrid). Imaginemos que, a fecha de la investigación, hubieran 4.500 personas colegiadas. La *muestra* (si se fija un margen de error de ±3% y un nivel de confianza de 2σ) estaría compuesta por 891 colegiados.

Si se tomase el número de colegiado como *número de identificación*, no habría que asignar a cada unidad de la población un número.

Para la extracción de los números se comenzaría por una fila o columna de una *tabla de números aleatorios*. En el Cuadro 5.5 figura un extracto de la *tabla de números aleatorios* adjunta en el Anexo. Del cuadro se deduce que la decisión tomada ha sido comenzar en la primera columna, escogiendo los cuatro últimos dígitos, y continuar hacia abajo hasta el final de la columna. Cuando ésta concluye, seguir en la siguiente columna, pero ahora tomando los cuatro primeros dígitos e, igualmente, hacia abajo. El procedimiento continúa hasta que, al menos, se alcanza el *tamaño de la muestra* fijado. Se decide extraer 60 colegiados más para posibles sustituciones de la *muestra* inicial.

Descartando los número superiores a la cifra de 4.500 (n.° total de colegiados), los sujetos seleccionados serían aquellos cuyo *número de identificación* (n.° de colegiado) se correspondiese con las cifras siguientes: 4357, 2111, 0801, 1767, 3740, 0924, 0364, 1247, 2147, 3364, 4460, 0621, 2815, 3666, 4410, 0054, 0935, 3602.

La *probabilidad de selección* dada a cada unidad de la población sería 0,198, al ser la *fracción de muestreo:*

$$\frac{n}{N} = \frac{891}{4.500} = 0{,}198$$

CUADRO 5.A. Extracto de una tabla de números aleatorios.

9	6754	1767	6	5	5659	4410	5	4	7361	3483	3
3	4357	8604	0	5	3364	7172	6	4	5690	6633	4
0	6318	3740	3	4	9927	5771	5	5	0423	6737	2
6	2111	5282	0	0	7243	7993	1	8	9292	8476	7
4	7534	0924	3	6	7879	0054	4	2	3410	1274	0
9	8614	7599	3	8	4460	6284	6	5	9844	1492	2
2	4856	0364	8	4	4898	0935	1	9	8795	1864	4
9	6887	1247	9	8	0621	6622	3	8	6085	7828	5
9	0801	2147	2	4	2815	7740	8	3	7390	7676	6
5	5165	7731	2	8	3666	3602	8	2	8420	7021	9

La exigencia de disponer de un listado de la población de interés, y su numeración, hace desaconsejable este procedimiento de selección muestral en poblaciones de elevada magnitud. En estos casos, habría que acudir a otro tipo de *diseño muestral* que suponga un menor coste de realización.

El Cuadro 5.5 resume las ventajas y los inconvenientes principales del *muestreo aleatorio simple*.

CUADRO 5.5. Ventajas e inconvenintes del muestreo aleatorio simple.

Ventajas	*Inconvenientes*
Facilidad de los cálculos estadísticos.	Requiere listar y enumerar a las unidades de la población.
Elevada probabilidad de lograr la equiparabilidad entre las características de la muestra y las correspondientes a la población.	Resulta monótono y arduo en muestras y poblaciones elevadas (sobre todo cuando se emplean procedimientos. manuales).
	La dispersión alcanzada en la muestra repercute negativamente en los costes de la investigación.

5.4.2. Muestreo aleatorio sistemático

Esta variedad de muestreo probabilístico es más practicada que la anterior. Exige igualmente la existencia de un listado de la población, pero difiere del *muestreo aleatorio simple* en dos aspectos fundamentales:

a) Sólo la primera unidad de la *muestra* se elige al azar (mediante una *tabla de números aleatorios,* por sorteo u otro procedimiento), con la condición de que el número seleccionado sea inferior al *coeficiente de elevación.*

El *coeficiente de elevación* se calcula dividiendo el *tamaño del universo* por el *tamaño de la muestra* (N/n). Expresa el número de veces que la *muestra* se halla contenida en el *universo.*

b) Los restantes elementos de la *muestra* se obtienen sumando, sucesivamente, el *coeficiente de elevación* (a partir del primer número elegido aleatoriamente), hasta completar, al menos, el *tamaño muestral.*

Si, previamente, no se ha extraído un excedente de unidades muestrales a considerar para las "sustituciones" (de los elementos de la *muestra),* en el momento de la *selección muestral* (o posteriormente), ha de calcularse un nuevo *coeficiente de elevación* que permita una nueva *selección sistemática* de las unidades muestrales no cubiertas en el *trabajo de campo* (por no estar localizables, por negación a cooperar u otra razón).

EJEMPLO DE MUESTREO ALEATORIO SISTEMÁTICO

Si la muestra del ejemplo anterior se eligiese de forma sistemática, se actuaría de la manera siguiente:

a) Se calcula el *intervalo de selección,* determinado por el *coeficiente de elevación*

$$\frac{N}{n} = \frac{4.500}{891} = 5,1$$

b) Se extrae al azar un número inferior a 5; por ejemplo, el 3. La unidad de la población a la que corresponda dicho número (en el marco muestral utilizado), constituye la primera unidad de la *muestra.*

c) Las demás unidades muestrales se eligen sumando, sistemáticamente, el *coeficiente de elevación* a los números que se seleccionen a partir del primer número extraído al azar. En este ejemplo, el 3 sería el primer número. Le seguirían el 8, el 13, el 18, 23, 28, 33..., hasta reunir, como mínimo, las 891 unidades que integran la *muestra* del estudio.

La *muestra* así extraída se asemejaría a la lograda mediante un *muestreo aleatorio simple*

En el Cuadro 5.6 figuran algunas de las ventajas y de los inconvenientes del *muestreo aleatorio sistemático.*

CUADRO 5.6. Ventajas e inconvenientes del muestreo aleatorio sistemático.

Ventajas	*Inconvenientes*
No precisa del uso continuo de una tabla de números aleatorios (u otro procedimiento de asignación aleatoria). No exige que el marco muestral sea un listado. Éste, en cambio, puede adoptar varias formas (fichas, papeletas,... hasta la propia presencia física de las unidades de la población).	Necesita del recuento constante de las unidades de la población. Antes de la selección muestral, hay que desordenar el marco muestral (si éste se encuentra ordenado de de acuerdo con algún criterio que favorezca la mayor representación de determinados segmentos de la población en la muestra).

5.4.3. Muestreo aleatorio estratificado

El *muestreo aleatorio estratificado* constituye una de las modalidades de muestreo más practicadas en la investigación social, cuando se dispone de información sobre características de la población de interés. Supone la clasificación de las unidades de población (contenidas en el *marco de muestreo* escogido), en un número reducido de grupos *(estratos),* en razón de su similaridad, dictada por las características observadas en el *marco muestral.* Con ello se persigue que cada *estrato* tenga representación en la *muestra* final.

Al igual que en el *muestreo por cuotas,* en el *estratificado* la *muestra* se distribuye en diferentes grupos de población, en función de los valores que presente en las variables elegidas para la *estratificación.* No obstante, ambas modalidades de muestreo difieren en la forma de elección de las unidades muestrales. Mientras que en el *muestreo por cuotas* intervienen criterios subjetivos, en el *estratificado* sólo el *azar.* En él, la extracción de las unidades finales de la *muestra* (en cada *estrato)* se hace siguiendo exclusivamente procedimientos aleatorios de selección muestral *(muestreo simple, sistemático, de rutas aleatorias).*

Lynn y Lievesley (1991) destacan cuatro principios básicos en el diseño de un esquema de *estratificación:*

a) *Elección de las variables de estratificación.* En general, la elección de las variables está condicionada a aquellas comprendidas en el *marco muestral* de referencia. De ellas se escogerán las más relacionadas con los objetivos de la investigación, con la finalidad de aumentar la efectividad de la *estratificación.*

   De las variables también se exige que sean *categóricas.* En caso contrario, habría que cambiar su nivel de medición; es decir, categorizarlas.

b) *Orden de las variables de estratificación.* Como la primera variable de estratificación es la que más discrimina, en el primer estadio de *estratificación* debería

elegirse la variable de mayor relevancia para la investigación; en el segundo estadio, la segunda variable más importante; y, así, consecutivamente, debido al poder decreciente de la *estratificación*.

c) *Número de variables de estratificación*. Para la mayoría de los objetivos de investigación, no se precisa diferenciar tres o cuatro estadios de *estratificación*. Tampoco se requiere la adopción de las mismas variables de estratificación para todas las muestras. Puede alcanzarse una mayor eficacia siguiendo un esquema de *estratificación* distinto para la variables incluidas en los diversos estadios de la *estratificación*.

d) *Tamaño de los estratos*. La *estratificación* resulta más adecuada, cuando se divide la población en grupos de igual tamaño. En cambio, se alcanzan escasos beneficios, cuando se utilizan variables bastante correlacionadas entre sí. En este caso, los estratos diferenciados variarían en su tamaño, aunque no en su naturaleza, en contra del propósito primordial de la *estratificación:* la consecución de grupos de población muy homogéneos entre sí y diferentes del resto.

En suma, si con la *estratificación* se persigue el logro de una mayor precisión en la estimación muestral, ésta se alcanzará cuando se cumplan dos condiciones esenciales:

a) Sean máximas las diferencias entre los estratos y mínimas dentro de cada estrato.

b) Las variables de estratificación se hallen relacionadas con los objetivos de la investigación (con las variables independientes y/o dependientes). De no ser así, se obtendría una precisión similar a la alcanzada sin la *estratificación*.

EJEMPLO DE ESTRATIFICACIÓN

En un estudio sobre la incidencia del fracaso escolar en la conducta delictiva, se elegirían como *variables de estratificación* cualquiera de las habitualmente asociadas a la causalidad de ambos fenómenos sociales: el fracaso escolar y la delincuencia.

Es fácil que el investigador disponga de información sobre las variables sexo y clase social. En caso afirmativo, podría estratificarse en función de ambas variables. Atendiendo a estas variables, la población quedaría, por ejemplo, dividida en seis *estratos*: varones de clase alta, varones de clase media, varones de clase baja, mujeres de clase alta, mujeres de clase media, mujeres de clase baja.

De esta forma, el investigador asegura la representación en la *muestra* de los estratos diferenciados. Lo que será de utilidad para la posterior realización de comparaciones entre los datos obtenidos de cada uno de los grupos señalados.

Las *variables de estratificación* más empleadas son las variables sexo y edad, debido a que ambas se hallan recogidas en la mayoría de los *marcos muestrales*. A ellas pueden añadirse otras variables, como la clase social, la ocupación, el nivel de instrucción; depende del objetivo fundamental de la investigación.

En estudios a nivel nacional (e internacional), suele estratificarse por ubicación geográfica: ámbito territorial (país, comunidad autónoma, provincia, municipio), tipo de hábitat (urbano, semiurbano, rural), o por tamaño de hábitat (número de habitantes).

Tras la clasificación de la *población* en *estratos,* se procede a *afijar la muestra* en cada *estrato*. Por *afijación* se entiende la distribución del *tamaño muestral* global entre los *estratos* diferenciados. Esta distribución puede cumplirse de tres maneras distintas: de forma *simple, proporcional* u *óptima*.

• *Afijación simple*

Se asigna el mismo *tamaño de la muestra* a cada *estrato*. Con ello se busca la igual representación de los estratos en la muestra global. Lo que favorecerá el logro de estimaciones igualmente significativas en todos los estratos.

Esta equidistribución del *tamaño muestral* conlleva, no obstante, un *inconveniente* importante: favorece a los estratos de menor volumen de población (al concedérseles un tamaño muestral superior al correspondiente a su volumen poblacional), respecto de los estratos de mayor magnitud.

Para la *afijación simple* se divide el *tamaño muestral* global entre el número de *estratos*. La cantidad así obtenida expresará las unidades a observar en cada uno de los estratos.

• *Afijación proporcional*

Como su nombre indica, en la *afijación proporcional* la distribución de la *muestra* se hace "proporcional" al peso relativo del *estrato* en el conjunto de la población. Por lo que, a los estratos que reúnan un mayor número de unidades de población les corresponderá un tamaño muestral superior al de aquellos que representen un porcentaje inferior en la población.

Este criterio de afijación es el más seguido. Para su cálculo, se multiplica la *proporción* que representa el *estrato* en la población por el *tamaño muestral* a afijar.

• *Afijación óptima*

Al peso relativo del *estrato* en la población, en la *afijación óptima* se añade la variabilidad (o heterogeneidad) del *estrato* respecto a la variable considerada en la estratificación. En conformidad con este último criterio de *afijación,* les corresponderá un tamaño muestral superior a los *estratos* de mayor heterogeneidad y peso poblacional.

La realización de esta variedad de *afijación* precisa del conocimiento previo de la *varianza poblacional* en cada *estrato,* respecto de las variables que se hayan empleado para la *estratificación.* Esta información es difícil de conocer (los *marcos muestrales* usuales no la facilitan), a menos que se haya analizado con anterioridad a la misma población (en un *estudio piloto ex profeso* o en investigaciones periódicas). Esto explica la escasa práctica de este tipo de *afijación,* pese a su calificativo de "óptima".

Para su cálculo se multiplica, en cada *estrato,* el *porcentaje* que representa el *estrato* en el conjunto de la población por la *varianza* (o *desviación típica)* correspondiente. Luego se suman todos los productos y se calcula la proporción que representa cada producto en el total. Por último, se multiplican estas proporciones por el tamaño de la muestra. De esta forma se obtiene el número de unidades de la población a observar en cada estrato.

EJEMPLO DE AFIJACIÓN DEL TAMAÑO MUESTRAL EN ESTRATOS

En una investigación sobre las actitudes delictivas de la población universitaria española, se decide la *estratificación* de la población universitaria por nivel académico, con la finalidad de garantizar la presencia en la *muestra* de los distintos niveles académicos. La *muestra* global está integrada por 2.500 unidades (con un *error* máximo de ± 2% y *nivel de confianza* de 2σ). Esta *muestra* se *afija* en los estratos siguiendo alguno de los criterios siguientes:

Nivel de estudios universitarios	*Porcentaje población*	*Varianza*	*Afijación*		
			Simple	*Proporcional*	*Optima*
Primer ciclo	45	1.900	833	1.125	970
Segundo ciclo	39	2.600	833	975	1.150
Tercer ciclo	16	2.100	833	400	380
			2.499	2.500	2.500

Afijación simple: $\frac{2500}{3} = 833$

De cada estrato se escoge el mismo número de unidades muestrales

Afijación proporcional:

$0{,}45 \cdot 2.500 = 1.125$
$0{,}39 \cdot 2.500 = 975$
$0{,}16 \cdot 2.500 = 400$

Al primer estrato le corresponderá un mayor tamaño muestral porque reúne un volumen de población más amplio. En cambio, al tercer estrato sólo se le asignan 400 unidades muestrales.

Afijación óptima:

(A) Porcentaje	(B) Varianza	(C) (A) . (B)	(D) Proporción (C)	(E) Afijación óptima*
45	1.900	85.500	0,388	970
39	2.600	101.400	0,460	1.150
16	2.100	33.600	0,152	380
100		220.500	1.000	2.500

(*) La columna (E) se obtiene multiplicando cada cifra que figura en la columna (D) por el tamaño muestral a fijar (2.500 unidades).

Como el segundo estrato presenta la *varianza poblacional* más elevada y, al mismo tiempo, es el segundo con mayor representación poblacional, éste resulta ser el más favorecido en la *afijación óptima*.

Por último, compárense las cifras obtenidas en las tres modalidades de *afijación*. Así podrán apreciarse las variaciones en la distribución de la *muestra* que producen los tres criterios de *afijación* existentes.

Las tres variedades de *afijación* pueden englobarse en dos amplias modalidades de *estratificación:* la *estratificación proporcional* y la *no proporcional.*

- *Estratificación proporcional*

La *estratificación* se hace de manera que garantice una probabilidad igual de selección para todos los estratos. Como difiere el conjunto de unidades que forman el estrato, los tamaños muestrales correspondientes a cada estrato también variarán en proporción similar.

- *Estratificación no proporcional*

La representación de los estratos en la *muestra* final no es proporcional a su peso en el conjunto de la población, al haberse dado una probabilidad desigual de selección en cada estrato (mediante la *afijación simple* o la *óptima).*

Esta "no proporcionalidad" puede deberse al deseo de analizar, con mayor detalle, unos estratos concretos, a los cuales les correspondería un tamaño muestral inferior, si se hubiese optado por la *estratificación proporcional;* o, simplemente, para propiciar la representatividad de las estimaciones muestrales en todos los estratos.

Uno de los *inconvenientes* fundamentales de la *estratificación no proporcional* es la necesidad de *ponderar la muestra,* si se desea la obtención de estimaciones muestrales para el conjunto de la población. Por el contrario, no se precisa de la *ponderación,* cuando sólo se realizan análisis individuales y/o comparativos entre los estratos.

Por *ponderación* se entiende el proceso de asignación de "pesos" a cada *estrato,* de manera que logre compensarse la desigual probabilidad de selección dada a cada unidad de población que compone el *estrato.* A tal fin, se comparan los datos muestrales con características de la población de interés publicadas en el último *censo de población, padrón de habitantes* u otro sondeo a cuyos datos se conceda una significativa *validez.* Por lo que, antes de *ponderar,* el investigador deberá comprobar la adecuación de los datos que toma como referente de las características poblacionales.

La *ponderación* puede efectuarse de varias formas. La más usual consiste en dividir el porcentaje que representa el estrato en el *universo* entre el porcentaje que representa en la *muestra.* De esta manera se obtienen los "pesos" o *coeficientes de ponderación.* Éstos se multiplicarán por cada una de las *estimaciones muestrales* de los estratos correspondientes para, así, poderse inferir el valor de los respectivos *parámetros poblacionales.*

En los ficheros de datos de la mayoría de los paquetes informático se fijan los *pesos* a asignar a cada subconjunto *(estrato)* de población. El programa automáticamente cumplimenta la *ponderación,* antes de proceder a la *tabulación* conjunta y posterior análisis de la información obtenida.

EJEMPLO DE PONDERACIÓN

Una institución alavesa encarga una encuesta para conocer los hábitos de consumo de droga entre la población (de ambos sexos) de 16 y más años de edad, por franjas de edad y zonas geográficas. El *tamaño de la muestra* se ha fijado en 1.600 entrevistas. Lo que supone un *error genérico de muestreo* del ± 2,5%, para un *nivel de confianza* del 95,5% (2σ) y un supuesto de *varianza* máxima (P = Q = 0,50).

Dada la distribución de la población alavesa en el territorio (en la que destaca la concentración en la capital de tres de cada cuatro habitantes), el equipo investigador decide realizar una *afijación* inicial *no proporcional* de la *muestra* total. Se asignan 900 entrevistas a la capital Vitoria (con 172.546 habitantes de 16 y más años de edad) y las 700 restantes distribuidas en el conjunto de los otros municipios (con una población global de 56.224 habitantes de 16 y más años).

El *objetivo* es obtener resultados con *errores de muestreo* similares en estas dos grandes zonas geográficas. Para devolver a los datos su proporcionalidad (y poder así analizar

conjuntamente la información recabada en cada estrato), se procede a la *ponderación de la muestra:*

Cuadrillas que dividen administrativamente Álava	*Población 16 y más años (*)*	*(A) Porcent. poblac.*	*Afijación propor. (**)*	*Afijación no propor. (***)*	*(B) Porcentaje muestra*	*Coeficiente ponderación (A/B)*
Añana	4.868	2,1	34	61	3,8	.553
Ayala	28.263	12,4	198	352	22,0	.564
Campezo	2.880	1,3	21	36	2,3	.565
La Guardia	8.066	3,5	56	100	6,2	.565
Salvatierra	7.434	3,2	51	92	5,7	.561
Vitoria	172.546	75,4	1.206	900	56,3	1.339
Zuya	4.713	2,1	34	59	3,7	.568
	228.770	100	1.600	1.600	100	

(*) *Fuente: Censo de Población 1991,* INE.

(**) Número de entrevistas que se habrían realizado, si se hubiese efectuado una *afijación proporcional.* Estas se obtienen multiplicando la proporción que representa cada *cuadrilla* (zona) en el conjunto de la población por el tamaño muestral (1.600).

(***) Número de entrevistas que en realidad se han efectuado: 900 en Vitoria y las 700 restantes distribuidas proporcionalmente entre el resto de las cuadrillas. Para ello se calcula nuevamente la proporción que representa cada cuadrilla, pero en el conjunto de las seis cuadrillas (eliminando a Vitoria), con una población total de 56.224 habitantes. Las proporciones así obtenidas se multiplican por el número de entrevistas a afijar en las seis cuadrillas (700). Por ejemplo, la primera cuadrilla representa ahora el 8,7% de la población total (exceptuando Vitoria). Multiplicando esa proporción por 700, se tendrían las 61 entrevistas a realizar; 27 más que si se hubiese hecho una *estratificación proporcional.*

En suma, si se hubiese *afijado proporcionalmente,* el *tamaño muestral* correspondiente a las cuadrillas (a excepción de Vitoria) habría sido bastante inferior al conseguido mediante la *estratificación no proporcional.* Mediante la *ponderación* (o equilibraje de la *muestra)* se devuelve a los datos su proporcionalidad, antes de proceder a la tabulación conjunta. Para ello se indican los *coeficientes de ponderación* (o *pesos)* en el fichero de datos y, automáticamente, se multiplica cada valor numérico obtenido por su correspondiente *coeficiente de ponderación.* De esta forma se obtienen las *estimaciones muestrales,* a partir del instrumento de medición aplicado en la investigación.

A modo de resumen, en el Cuadro 5.7 se indican algunas de las ventajas e inconvenientes que lleva consigo la aplicación de un *muestreo aleatorio estratificado.*

CUADRO 5.7. Ventajas e inconvenientes del muestreo aleatorio estratificado.

Ventajas	*Inconvenientes*
Supone un menor error muestral y, por tanto, una mayor precisión de la estimación muestral.	Precisa más información del marco muestral que el muestreo aleatorio simple (para identificar a la población de cada estrato). Ello puede resultar costoso.
Asegura la representación de las variables de estratificación (y de las variables relacionadas con ellas).	Lleva consigo cálculos estadísticos complejos.
Pueden emplearse procedimientos muestrales variados en los distintos estratos.	
Facilita la organización del trabajo de campo.	

5.4.4. Muestreo aleatorio por conglomerados

El *muestreo por conglomerados* también representa un procedimiento de selección aleatoria de un conjunto de individuos (referidos ahora como *conglomerados).*

Con el *muestreo aleatorio estratificado* comparte la característica básica de seccionar la población total en grupos, como fase previa a la extracción muestral. Si bien difiere de él en varios aspectos importantes:

a) En el *muestreo estratificado* se busca la homogeneidad dentro del estrato y la heterogeneidad entre los estratos. En el *muestreo por conglomerados* es a la inversa: el *error muestral* disminuye conforme aumenta la heterogeneidad dentro del grupo *(conglomerado).* Ello se debe a la necesidad de que cada *conglomerado* constituya una representación, lo más ajustada posible, de la variedad de componentes del universo.

b) En el *muestreo estratificado* se selecciona aleatoriamente una *muestra* para cada *estrato*. En el *muestreo por conglomerados* lo que se extrae es una muestra aleatoria de *conglomerados*. Sus integrantes formarán la *muestra.*

c) En el *muestreo estratificado,* la unidad de muestreo es el individuo. En cambio, en el *muestreo por conglomerados* es el *conglomerado* (o conjunto de individuos).

Los *conglomerados* pueden ser las áreas geográficas que dividen a la población que se analiza (país, comunidad autónoma, municipio, distrito, áreas censales, vivien-

das); pero, también, organizaciones u instituciones (colegios, hospitales, tribunales, centros penitenciarios).

Sudman (1976) señala los siguientes aspectos a considerar en la elección de los *conglomerados:*

a) Los *conglomerados* han de estar bien definidos y delimitados. Cada unidad de la población sólo puede pertenecer a un único *conglomerado.*
b) El número de elementos que componen el *conglomerado* ha de ser conocido previamente (aunque sea de manera aproximada).
c) Los *conglomerados* elegidos han de ser pocos, si realmente quieren reducirse los costes de la investigación.
d) Los *conglomerados* deberían escogerse de manera que se consiguiera disminuir el aumento en el *error muestral,* generado por la agrupación (o aconglomeración).
e) Los *conglomerados* no tienen por qué hallarse idénticamente definidos en todos los lugares.
Cuando se muestrean individuos u hogares en áreas urbanas, los *conglomerados* suelen ser bloques o conjuntos de bloques. En cambio, en las áreas rurales, los *conglomerados* serán segmentos geográficos limitados por carreteras y fronteras naturales (como ríos y lagos).

Si, a partir de una *muestra por conglomerados,* se extrae una nueva muestra, con referencia a cada uno de los *conglomerados* previamente elegidos, y así, sucesivamente, se está ante un *diseño muestral* muy habitual en la investigación social: el *muestreo polietápico por conglomerados.*

El *muestreo polietápico (o polifásico) por conglomerados* representa una extensión del *muestreo por conglomerados.* En él la unidad de muestreo final no son los *conglomerados,* sino subdivisiones de estos. Por lo que no se toman cada uno de los integrantes de los *conglomerados* elegidos aleatoriamente, sino sólo a una parte de ellos, escogidos también de forma aleatoria. Ello supone muestrear –como su nombre indica– en distintos niveles, implicando varios procedimientos de selección muestral (como la *estratificación).*

La modalidad de *muestreo polietápico por conglomerados* más sencilla implica la extracción muestral en dos fases (Henry, 1990):

a) En la *primera fase,* se seleccionan las agrupaciones de los miembros de la población de estudio (conocidas como las *unidades de muestreo primarias),* que son análogas a los *conglomerados.*
b) En la *segunda fase,* se eligen aleatoriamente los miembros de la población a observar, de las *unidades de muestreo primarias* previamente seleccionadas.

Pero, lo usual es que el número de estadios comprendidos en el *muestreo polietápico* sean tres o cuatro (Lynn y Lievesley, 1991). En el caso más general (en muestras

nacionales), se recurre a *muestreos polietápicos, estratificados, por conglomerados:* primero, se *afija* la *muestra* por *estratos;* después, se extraen (de forma aleatoria proporcional) los municipios, las secciones estadísticas; y, por último, los hogares en los que se realizarán las entrevistas (Osuna, 1989).

EJEMPLO DE MUESTREO ALEATORIO POR CONGLOMERADOS

Se quiere encuestar (mediante entrevista personal) a 1.300 profesores universitarios españoles, con el propósito de conocer su opinión sobre los nuevos planes de estudio.

Como la encuesta es de ámbito nacional, la selección de la *muestra* siguiendo el procedimiento *aleatorio simple* elevaría considerablemente los costes de la investigación. A la dificultad de encontrar un *marco muestral* que incluyera a todos los profesores universitarios españoles (tanto de universidades públicas como privadas), se sumaría el aumento de los costes por desplazamiento de los entrevistadores a puntos dispersos del país.

Lo más fácil sería acceder a un listado que incluyese al conjunto de las universidades españolas (públicas y privadas) para, a continuación proceder a la extracción *muestral polietápica por conglomerados.* Ésta podría consistir en:

a) Listar todas las universidades españolas, tanto públicas como privadas. Este listado constituiría el *marco de muestreo* inicial de la investigación.

En la selección de la muestra nacional de universidades podría *estratificarse* por *tipo de universidad* (pública y privada; y dentro de esta última, diferenciándose por confesionales y no confesionales), y por *ubicación geográfica* (Comunidad Autónoma). Ello aseguraría la representación de las *unidades de muestreo primarias* (las universidades) en conformidad con las *variables de estratificación* elegidas: dispersión geográfica y tipo de universidad.

También conviene –para alcanzar una mayor *precisión*– que las *unidades primarias de muestreo* elegidas sean bastante heterogéneas respecto a las características que se analizan. De esta forma se asegura que la variedad del universo de estudio quede reflejada en la *muestra.*

Asimismo, se recomienda que –como regla– no proceda más del 7% de la *muestra* total de una sola *unidad de muestreo primaria* (Lynn y Lievesley, 1991). Lo que interesa es que haya un número elevado de *conglomerados* en la primera fase del *diseño muestral.* Esto favorece la heterogeneidad de la *muestra,* con la consiguiente reducción del *error muestral.*

b) Elegir (de las universidades seleccionadas) una muestra de facultades mediante alguno de los procedimientos de selección aleatoria *(simple, sistemática, estratificada).*

c) Para cada una de las facultades escogidas, extraer una muestra aleatoria de profesores. A tal fin se acude a un nuevo *marco muestral:* el listado de los profesores pertenecientes a las facultades elegidas. Éstos podrían, a su vez, muestrearse conforme a los criterios de categoría profesional (catedrático, titular, profesor asociado, ayudante), dedicación (a tiempo completo o parcial), y antigüedad docente, siguiendo un *muestreo aleatorio estratificado* o *por cuotas.*

Si se opta por el *estratificado,* la elección de los profesores cuyas características se ajusten a los *estratos* diferenciados será totalmente aleatoria *(muestreo simple, sistemático).* En cambio, si se decide el de *cuotas,* la elección de los profesores será arbitraria, con la única condición de que compartan las características fijadas en las *cuotas.*

El *diseño muestral* que resulte de la conjunción de estas fases se corresponderá con la precisión que el investigador quiera dar a las *estimaciones muestrales.* Esto tiene una repercusión directa en los costes de la investigación.

Así, por ejemplo, en la elección del número de facultades y de profesores a entrevistar en cada una de ellas, las opciones posibles son diversas. Dependerá del presupuesto destinado a la realización del *trabajo de campo* y de la heterogeneidad de los *conglomerados,* principalmente. El investigador tendrá que decidir si aumentar el número de facultades, disminuyendo el número de profesores a entrevistar en cada una de ellas, o proceder a la inversa (reducir el número de facultades, incrementando el número de profesores a encuestar). La primera opción supone aumentar –para un mismo *tamaño muestral* (aproximadamente, 1.300 unidades)– los costes (en tiempo y dinero) de la investigación, respecto de la segunda opción que resulte más económica.

A continuación se exponen algunas de las opciones posibles:

N° facultades	*N° profesores*
5	260
10	130
15	87
20	65
30	43
40	33

Si fuesen 5 las facultades elegidas (entrevistándose a 260 profesores, en cada una de ellas), los costes del trabajo de campo serían inferiores; pero también sería mayor el *error muestral* que si se eligiesen 40 facultades, y de ellas se entrevistase a 33 profesores.

En general, se recomienda aumentar el número de *conglomerados* (en este caso de facultades) con preferencia a elevar el número de individuos a observar en cada uno de ellos (los profesores). La razón está en la probable homogeneidad de los *conglomerados* elegidos.

Cuando éstos son bastante homogéneos, no se precisa añadir más elementos del *conglomerado* a la *muestra,* ya que se obtendría una información redundante, al ser similares las características de las unidades que forman el *conglomerado.* En este caso (cuando los *conglomerados* son homogéneos), se aconseja ampliar el número de *conglomerados* para, de esta forma, abarcar una mayor variedad de la población de interés. Por el contrario, si los *conglomerados* fuesen heterogéneos, la mejor opción (entendida como reducción del *error muestral)* sería la opuesta: reducir el número de *conglomerados,* aumentando las unidades a observar en cada uno de ellos.

La probabilidad de selección de cada uno de los integrantes de los *conglomerados* será igual al número de unidades a elegir de cada *conglomerado* dividido por el número de unidades contenidas en el *conglomerado.*

En suma, el *muestreo aleatorio por conglomerados* se muestra de especial interés cuando:

a) Resulte difícil compilar una lista exhaustiva de todos los componentes de la población. Lo que imposibilitaría la práctica de otra variedad de *muestreo probabilístico*.
b) Se quiera reducir la duración y los costes económicos del trabajo de campo en la investigación.
c) Se realicen estudios de ámbito nacional o internacional, que supongan una considerable dispersión de la *muestra*.

El Cuadro 5.8 resume ventajas e inconvenientes principales del *muestreo aleatorio por conglomerados*.

CUADRO 5.8. Ventajas e inconvenientes del muestreo aleatorio por conglomerados.

Ventajas	*Inconvenientes*
No exige un listado de toda la población de interés; sólo de las unidades del conglomerado.	Mayor error muestral (y, por tanto menor precisión de las estimaciones muestrales) porque los conglomera do suelen ser muy homogéneos *.
Al concentrar el trabajo de campo en un número limitado de puntos de muestreo, disminuyen los costes de la investigación.	Requiere cálculos estadísticos complejos en la estimación del error muestral, principalmente.

* El error muestral puede reducirse aumentando el número de *conglomerados*.

5.4.5. *Muestreo por cuotas*

Una de las modalidades de muestreo más populares es el *muestreo por cuotas*. Esta constituye una variedad de *muestreo no probabilístico* que parte, igualmente, de la segmentación de la población de interés en grupos, a partir de variables sociodemográficas relacionadas con los objetivos de la investigación. Por lo que, su puesta en práctica conlleva, también, la elaboración de una *matriz* con las características básicas de la población que se analiza (proporciones de población diferenciadas por sexo y edad, nivel de instrucción, clase social). Esta información suele obtenerse del último *censo de población, padrón de habitantes* u otra fuente estadística similar.

El propósito es seleccionar una *muestra* que se ajuste a la distribución de las características fundamentales de la población. Si el objetivo del estudio fuese, por *ejemplo,* comprobar la influencia del nivel educativo de la personas en sus actividades de ocio, la población debería, al menos, *estratificarse* por nivel educativo. Ello ga-

rantizaría que en la *muestra* se encuentren representados los distintos grupos de población (diferenciados por nivel de instrucción). Lo que ayudaría a la realización de comparaciones entre ellos.

Además de los objetivos del estudio, en la elección de las variables intervienen otros factores: la precisión que el investigador desee para su indagación, junto a la accesibilidad de las variables elegidas. Ésta dependerá de la facilidad de su obtención en el *marco de muestreo* elegido pero, también, de su practicabilidad (si el entrevistador puede acceder fácilmente) a los grupos de población definidos por las variables escogidas.

Las *cuotas* más habituales son las determinadas por la conjunción de las variables sexo y edad, en consonancia con su mayor accesibilidad. La mayoría de los *marcos muestrales* contienen ambas variables. A esto se suma la relativa facilidad (para el entrevistador) de localizar visualmente a personas que pertenezcan a los distintos grupos de sexo y edad.

Lo mismo no cabe decir de otras variables, como la clase social, por ejemplo. Esta variable se enfrenta –de acuerdo con Moser y Kalton (1989)– a dos importantes trabas que dificultan su uso en el establecimiento de las *cuotas:*

a) La carencia de fuentes estadísticas fiables para la delimitación de las *cuotas*. A ello contribuye la propia definición de la variable clase social, que supone la combinación de factores objetivos (ocupación, renta) y subjetivos (apariencia física, forma de vestir, de hablar).
b) La vaguedad de la definición de la variable clase social, que provoca discrecionalidad en el juicio subjetivo del entrevistador, con la consiguiente introducción de sesgos importantes en la investigación.

Una vez confeccionada la matriz, se calculan –siguiendo un proceso similar al *muestreo estratificado proporcional*– las proporciones relativas (las *cuotas)* para cada celdilla de la *matriz,* a partir de la proporción que representa cada categoría de las variables seleccionadas en la población total. Con ello se busca la equiparación de la *muestra* con la *población* de interés.

Después de la delimitación de las *cuotas,* se proporciona a cada entrevistador su asignación correspondiente: el perfil y el número de personas a entrevistar en cada *cuota*.

EJEMPLO DE ASIGNACIÓN DE CUOTAS

En un *diseño muestral* en el que se seccione la población conforme a las variables sexo, edad y clase social, primero se delimitan las *cuotas*, en consonancia con el peso de las distintas categorías de las variables en el conjunto de la población. Después, se asigna a ca-

da entrevistador un número total de entrevistas a realizar, acordes con las *cuotas* fijadas. La asignación podría ser la siguiente:

		Clase social			
		Alta	Media	Baja	Total
Varones	18 – 30	1	2	1	4
	31 – 45	1	5	2	8
	46 – 65	2	6	1	9
	Más de 65	–	1	–	1
Mujeres	18 – 30	–	2	1	3
	31 – 45	1	6	2	9
	46 – 65	2	5	2	9
	Más de 65	–	1	1	2
	Total	7	28	10	45

Lo que distingue al *muestreo por cuotas,* respecto del *estratificado,* es la libertad que se da al entrevistador para la elección de las unidades finales de la población a entrevistar. Aunque el azar intervenga en la fases iniciales del *diseño muestral* (en la elección de áreas o zonas geográficas, por ejemplo), la selección de los elementos concretos de la población es totalmente arbitraria. Es el propio entrevistador quien elije al entrevistado, en cualquier momento y lugar (en una calle comercial, a la salida del metro, en la parada del autobús, en un mercado, en un parque). La única condición que se le impone es que la persona se ajuste a las *cuotas* fijadas por el equipo investigador.

Este margen de libertad que se concede al entrevistador representa la principal debilidad de esta modalidad de muestreo:

a) El entrevistador es libre de entrevistar a quien quiera o pueda. Lo que puede introducir *sesgos* importantes en el proceso de selección muestral, no siempre detectables por el equipo investigador. Por *ejemplo,* puede haber escogido a personas de su propio entorno sociofamiliar, a aquellas que haya estimado más propensas a participar en la investigación, o simplemente a las más accesibles.

También es fácil que, dentro de una misma *cuota,* se escoja a unos individuos con preferencia a otros. Por *ejemplo,* si se controla por edad, es probable que se entreviste más a personas de edades intermedias en los intervalos diferenciados, y no a aquellas situadas en los límites del intervalo.

Asimismo, el entrevistador puede ubicar a los sujetos en *cuotas* diferentes a las que realmente pertenecen, en aquéllas donde se precisen casos (para completarlas). Esta última picaresca también resulta difícil de controlar.

b) El desconocimiento de las probabilidades de selección no permite evitar los errores generados por ponderaciones incorrectas en el proceso de estimación. Tampoco permite la estimación de los *errores de muestreo*.

Todo ello repercute en el principal *inconveniente* de este tipo de muestreo: la *muestra* finalmente obtenida puede no ser representativa de la población que se analiza, aunque la *muestra* diseñada coincida con la distribución de la *población* en los controles de *cuotas* fijados. Como Azorín y Sánchez Crespo (1986: 21) afirman, "puede ocurrir que la muestra represente muy bien a la población para ciertas variables, y no la represente en absoluto en cuanto opiniones".

Para solventar los *sesgos* inherentes en el *muestreo por cuotas,* éste suele complementarse con el *muestreo de rutas aleatorias:* para cada entrevistador se fija un *itinerario aleatorio* (definido en todos sus detalles), indicándole en qué puntos concretos ha de realizar cada entrevista. Estas indicaciones limitan la arbitrariedad del entrevistador.

En el Cuadro 5.9 se exponen ventajas e inconvenientes generalmente atribuidos al *muestreo por cuotas*.

CUADRO 5.9. Ventajas e inconvenientes del muestreo por cuotas.

Ventajas	*Inconvenientes*
Resulta más económico (en tiempo y dinero) que los muestreos probabilísticos.	Supone un mayor error muestral que los diseños probabilísticos.
Fácil de administrar.	Inexistencia de algún método válido para calcular el error típico (al no ser un muestreo probabilístico).
No precisa de un listado de la población.	Limites en la representatividad de la muestra para las características no especificadas en los controles de cuotas.
	Dificultad para controlar el trabajo de campo.

5.4.6. Muestreo de rutas aleatorias

El *muestreo de rutas aleatorias* también suele estar presente en el estadio final de un *diseño muestral* (en la selección de las unidades últimas de muestreo), comple-

mentando tanto a *muestreos no probabilísticos* (como el de *cuotas)* como a *probabilísticos* (el *muestreo de conglomerados* y/o el *estratificado).*

Se denomina *muestreo de rutas* porque se establece la "ruta" o *itinerario* que el entrevistador ha de seguir en la selección de las unidades muestrales. Las *rutas* se eligen de forma "aleatoria", sobre un mapa del municipio concreto donde han de realizarse las entrevistas. En el mapa se indican múltiples puntos de partida o de inicio de rutas posibles. Normalmente, el *punto de partida* corresponde a un edificio determinado por una dirección (por ejemplo, la calle Serrano n.º 22). Pero, en entidades de población pequeñas (núcleos rurales, por ejemplo), el *punto de inicio de la ruta* suele corresponder a edificios centrales, como el ayuntamiento, la iglesia o la plaza del pueblo.

Una vez que se ha elegido aleatoriamente el comienzo de la *ruta,* el entrevistador deberá tomar una dirección u otra, siguiendo las normas fijadas por el equipo investigador. En estas *normas* se indican:

a) Los *giros* que han de realizarse: a la izquierda o a la derecha.
b) Los *edificios* en los que debe entrarse: como los terminados en una cifra específica (por ejemplo, los acabados en 2).
c) Si el *edificio* comprende más de una *vivienda,* se especifica la elección de la *escalera* (si hay más de una); de la *planta;* de la *vivienda* en la planta (siguiendo una *tabla aleatoria);* y, por último, de la *persona* a entrevistar en cada *vivienda* (de acuerdo con las *cuotas* fijadas).

En las *hojas de ruta* el entrevistador registra las actuaciones seguidas hasta la localización de los sujetos de la *muestra,* así como cualquier incidencia de utilidad para la supervisión del *trabajo de campo.*

Este procedimiento de selección muestral por *rutas aleatorias* presenta la gran *desventaja* de no garantizar que todas las unidades de la población tengan la misma probabilidad de ser elegidas, aunque la designación de las *rutas* sea *aleatoria.* Además, existe el peligro de sobrerepresentación de las personas que se encuentran, con mayor frecuencia, en las viviendas (amas de casa, jubilados, personas en paro).

Para obviar dicha desventaja se aconseja complementar el *muestreo por rutas* con el *muestreo por cuotas.* También se reitera –véase Arber (1993)– la necesidad de insistir, al menos tres veces, en aquellas viviendas donde previamente no se haya obtenido respuesta.

5.4.7. Muestreo estratégico

El *muestreo estratégico* o *de conveniencia* responde a una modalidad de *muestreo no probabilístico,* en el que la selección de las unidades muestrales responde a criterios subjetivos, acordes con los objetivos de la investigación. Por lo que comparte las ventajas y los inconvenientes básicos de cualquier *muestreo no probabilístico* (Cuadro 5.10):

CUADRO 5.10. Ventajas e inconvenientes del muestreo estratégico.

Ventajas	*Inconvenientes*
Simplicidad y economía del diseño muestral.	Imposibilidad de estimar el error típico.
Fácil de administrar.	Dificultad en la generalización de los resultados de la investigación más allá de los casos analizados, por la introducción de sesgos en la elección de la muestra.
No precisa de un listado de la población.	

Esta variedad de *muestreo no probabilístico* es habitual en *estudios cualitativos,* no interesados en la generalización estadística. Pero, también es frecuente en los *experimentos* realizados con personas que se ofrecen voluntarias; e, igualmente, en *estudios piloto*, en los que se comprueba la validez del instrumento de medición elaborado (un cuestionario, un guión de entrevista). Estos estudios no siempre precisan de la selección rigurosa de las unidades muestrales. Basta con escoger (y no necesariamente de forma aleatoria) a individuos que compartan las características de la *muestra* diseñada para la investigación definitiva.

EJEMPLO DE MUESTREO ESTRATÉGICO

Un *diseño muestral* viable para el estudio sobre las trayectorias laborales de los jóvenes madrileños, mediante *entrevistas en profundidad,* sería *estratégico,* si la selección de los jóvenes madrileños se hiciera de forma no aleatoria, sino consecuente con criterios marcados por el investigador.

El investigador podría decidir que, para cumplir los objetivos de su estudio, basta con seleccionar un número reducido de jóvenes que constituyan casos "representativos" de la variedad juvenil actual (en sus comportamientos y características sociodemográficas). Para ello opta por delimitar la selección muestral a dos distritos polares de Madrid, como son los distritos de Salamanca y de Carabanchel. De ellos escoge, de forma deliberada, casos que juzga que "representan" al conjunto de la población cuyas características analiza.

El procedimiento de selección muestral concluye cuando se llega a la *"saturación teórica":* cuando la información comienza a ser redundante, no aportando ningún nuevo punto de vista analítico. Por lo que se decide la conclusión de la recogida de información.

5.4.8. Muestreo de "bola de nieve"

Esta última variedad de *muestreo no probabilístico* difiere del anterior en que las unidades muestrales van escogiéndose, sucesivamente, a partir de las referencias aportadas por los sujetos a los que ya se ha accedido. Como, a su vez, los nuevos casos identifican a otros individuos en su misma situación, la *muestra* va aumentando, como una "bola de nieve", hasta que el investigador decida cortar.

Este tipo de muestreo es de gran de utilidad, cuando se carece de un *marco de muestreo* que recoja a la población de interés. Especialmente, en poblaciones que, por sus propias características -generalmente población "marginal"-, son difíciles de identificar y de localizar. En estos casos, la localización de uno de los integrantes de la *población* se convierte en doblemente valiosa, al suponer la identificación de otras personas pertenecientes a la misma *población*. Para ello se pregunta a los sujetos ya "descubiertos", si conocen a otros en su misma situación.

Este procedimiento muestral comparte las ventajas y los inconvenientes comunes a los *diseños muestrales no probabilísticos*.

5.5. Ejemplos de diseños muestrales tomados de investigaciones reales

Como ilustración de lo expuesto en este capítulo, a continuación se detallan dos *diseños muestrales* correspondientes a dos encuestas nacionales: la Encuesta Nacional de Juventud 1988, dirigida por Zárraga, y el 2.° Informe de la Universidad Complutense, dirigido por A. de Miguel. Ambos *diseños muestrales* resultan de interés por la conjugación que hacen de varios procedimientos de selección muestral.

- Informe Juventud en España, 1988, *dirigido por José Luis de Zárraga Moreno (Ministerio de Asuntos Sociales, Instituto de la Juventud, 1989).*

La descripción del diseño de la muestra figura en el anexo de la investigación. En él se detallan los aspectos siguientes:

Universo de estudio

a) *Definición:* población de 15 a 29 años, residente en el territorio español (de la península Ibérica e islas Baleares), en hogares familiares situados en núcleos de población de distinto tamaño. Por lo que, se exceptúa a los jóvenes que residen en Canarias, Ceuta y Melilla, junto con aquellos que viven fuera del hogar familiar.

b) *Universo de referencia.* Se toma como universo de referencia el expuesto en el *Padrón municipal de 1986,* al ser ésta la fuente estadística más próxima al momento de la investigación (1988). El último *censo de población* era de fecha de 1981.

A efectos de distribución de la población en entidades dentro de cada municipio, se toman los datos del *Nomenclator* del *Censo de población de 1981.*

Para la fijación de las *cuotas* de sexo-edad-estado civil, el referente es *La encuesta de población activa,* publicada por el INE.

c) *Estratificación del universo.* La *población* se estratifica por región y tipo de hábitat. Para la *estratificación* (por región) se aplica la actual división político-administrativa del país en comunidades autónomas. Para la formación de los *estratos* según el tipo de hábitat, se aplica el criterio del tamaño del municipio, distinguiéndose las ocho categorías siguientes:

1) Menos de 2.000 habitantes.
2) De 2.000 a 5.000 habitantes.
3) De 5.001 a 10.000 habitantes.
4) De 10.001 a 50.000 habitantes.
5) De 50.001 a 100.000 habitantes.
6) De 100.001 a 250.000 habitantes.
7) De 250.001 a 1.000.000 habitantes.
8) Más de 1.000.000 habitantes.

El cruce de ambas variables (comunidad autónoma y tamaño del municipio) genera 128 estratos. De ellos, 26 quedan vacíos, por no existir en esa comunidad autónoma algún municipio del tamaño correspondiente.

Tamaño de la muestra

Considerados los objetivos del estudio, se estima aceptable un *error de muestreo* máximo de ±1,5% (con un *nivel de confianza* del 95,5%) en los datos a nivel nacional, para lo cual se precisa una *muestra* de 4.300 individuos, al menos.

Asimismo, se establece un *tamaño mínimo* de 625 entrevistas –que proporciona un *error de muestreo* máximo de ±4%– para las comunidades mayores y aquellas que presentan una diferenciación cultural superior respecto al conjunto nacional, con la finalidad de facilitar análisis específicos en cada una de ellas. Este criterio se aplica a las submuestras de las CC.AA. de Andalucía, Madrid, Cataluña, Galicia y País Vasco.

Por otra parte, también se adopta el criterio de ampliar las submuestras de solteros y de casados para tener datos significativos en ambos subconjuntos de población (de muy diferentes modos de vida).

Con estos criterios se diseña una *muestra básica* –con ampliaciones regionales– formada por 4.320 jóvenes, distribuidos por comunidades autónomas y tipos de hábitat como figura en el Cuadro 5.11. La asignación de entrevistas se hace de forma proporcional *(afijación proporcional)* al tamaño del *estrato* en la comunidad autónoma, respecto del total de España (según la fuente estadística consultada).

CUADRO 5.11. Distribución de la muestra base.

		Distribución de las entrevistas según tamaño del municipio							
	N.º entrevistas	*Menos 2.000*	*2.000 5.000*	*5.001 10.000*	*10.001 50.000*	*50.001 100.000*	*100.001 250.000*	*250.001 1 millón*	*Más de 1 millón*
Galicia	262 + 298	11	103	138	137	78	45	48	–
Asturias	109	2	7	9	30	20	17	24	–
Cantabria	52	6	9	4	9	6	18	–	–
País Vasco	223 + 337	29	25	42	149	72	127	116	–
Navarra	52	11	12	4	7	–	18	–	–
Aragón	112	30	9	6	16	–	–	51	–
Baleares	61	1	7	8	18	–	–	27	–
Cataluña	574	42	40	31	114	55	86	28	178
Castilla-León	260	99	29	14	23	23	42	30	–
Rioja	25	6	3	3	3	–	10	–	–
Extremadura	110	26	25	20	22	6	11	–	–
Madrid	474 + 85	10	8	8	29	34	76	–	395
Castilla-La Mancha	163	42	28	30	31	21	11	–	–
C.Valencia	359	27	29	32	123	22	27	99	–
Andalucía	667	37	80	77	186	53	61	173	–
Murcia	96	1	1	10	32	6	17	29	–
Total	3.600 + 720	380	415	436	929	396	566	625	573
	4.320								

A esta *muestra básica* se agrega una *ampliación de muestra:* por cada sujeto casado seleccionado en la *muestra básica* se elige (adicionalmente) un sujeto soltero del mismo sexo y grupo de edad. De este modo, se agregan 922 entrevistas adicionales a la *muestra básica*. La *muestra total* está, por tanto, formada por 5.242 jóvenes, distribución por sexo, edad y estado civil como se expone en el Cuadro 5.12.

CUADRO 5.12. Distribución de la muestra total por sexo, edad y estado civil.

	Varones			*Mujeres*		
	Total	Solteros	Casados	Total	Solteras	Casadas
15 - 19	747	746	1	746	757	7
20 - 24	795	733	62	859	715	144
25 - 29	1.001	705	296	1.076	664	412
Total	2.543	2.184	359	2.699	2.136	563

(Los desequilibrios introducidos en la muestra se corrigen durante el proceso de análisis, mediante *coeficientes de ponderación* que restituyen a los subconjuntos muestrales su peso correspondiente en la distribución del universo).

Selección de los puntos de muestreo

La primera etapa de la selección de la muestra consiste en la elección de los municipios dentro de cada estracto. En total se seleccionan 262 municipios, distribuidos por región y tipo de hátitat del modo siguiente (Cuadro 5.13):

CUADRO 5.13. Distribución geográfica de los puntos de muetreo.

	Distribución de las entrevistas según tamaño del municipio								
	Menos 2.000	*2.000 5.000*	*5.001 10.000*	*10.001 50.000*	*50.001 100.000*	*100.001 250.000*	*250.001 1 millón*	*Más de 1 millón*	*Total*
Galicia	1	12	9	9	4	1	1	–	37
Asturias	–	1	1	2	1	1	1	–	7
Cantabria	1	2	–	1	1	1	–	–	6
País Vasco	4	3	3	10	4	3	1	–	28
Navarra	1	2	–	1	–	1	–	–	5
Aragón	4	1	1	1	–	–	1	–	8
Baleares	–	1	1	1	–	–	1	–	4
Cataluña	6	5	2	8	3	4	1	–	30
Castilla y León	12	4	1	2	1	2	1	–	23
Rioja	1	–	1	–	–	1	–	–	3
Extremadura	3	3	2	2	1	1	–	–	12
Madrid	1	1	1	2	2	4	–	–	12
Castilla-La Mancha	6	4	2	2	1	1	–	–	16
Com. Valenciana	4	4	2	8	1	2	2	–	23
Andalucía	5	10	5	12	3	3	4	–	42
Murcia	–	–	1	2	1	1	1	–	6
Total	49	53	32	63	23	26	14	2	262

(La selección de los municipios dentro del estrato se hace aleatoriamente, asignando a cada municipio una probabilidad proporcional a su población.)

La segunda etapa consiste en la selección de núcleos de población dentro de cada municipio. Para esta selección se siguen los criterios siguientes:

a) En cada municipio menor de 50.000 habitantes se elige al azar una entidad de población, con probabilidad proporcional al número de habitantes registrados en el *Nomenclator* del *Censo de 1981*.

b) En los municipios mayores de 50.000 habitantes se calcula la proporción de la población que reside en la entidad principal y en el conjunto de las restantes entidades. Si, repartiendo el número de entrevistas asignadas al municipio (en proporción directa a esas poblaciones), corresponde al conjunto de las entidades menores de seis o más entrevistas de la *muestra básica,* entonces se selecciona aleatoriamente una de dichas entidades menores, y a ésta se le asignan las entrevistas. En caso contrario, todas las entrevistas asignadas al municipio se realizan en la entidad principal.

Proceso de selección de individuos

En una primera fase, la elección de los entrevistados se hace aplicando un *muestreo polietápico,* con selección de hogares por el procedimiento de *itinerarios aleatorios* y de individuos (dentro del hogar) mediante *números aleatorios.* En cada estrato muestral se aplica además un sistema de *cuotas* de edad y sexo para corregir las desviaciones que pudieran producirse por la incidencia de las negativas a participar en la investigación y las ausencias en los domicilios.

En una segunda fase del *trabajo de campo,* la correspondiente a la ampliación de individuos casados, se emplea un sistema de selección por *cuotas* de edad, sexo y estado, dentro de zonas elegidas aleatoriamente en las entidades de población de la *muestra básica.*

Estimación de los errores de muestreo

Con el procedimiento de selección muestral utilizado no es posible un cálculo exacto de los *errores muestrales.* En estos casos, suele tomarse como referencia los *errores* que corresponderían a un *muestreo aleatorio simple,* ya que, si bien la selección "en racimos" –como la del procedimiento de "*itinerarios aleatorios*"– incrementa esos errores, la *estratificación de la muestra* los disminuye.

- *Diseño muestral realizado por Valles Martínez, M. y Cea D'Ancona, M.ª A., para el 2.º Informe de la Universidad Complutense: De Miguel, A. (1994):* La Sociedad Española 1993-1994, *Madrid, Alianza.*

Para la realización de este *informe* se llevaron a cabo dos operaciones de encuesta: una dirigida a la población española en su conjunto, y otra a las personas de más de 64 años (mayores) que viven en residencias para la tercera edad. La primera, a personas de distintos grupos de edad y situación económica (jóvenes, adultos activos, amas de casa, y mayores). La segunda constituye, en cambio, una *muestra estratégica,* hecha con el propósito de contrastar la opinión de los ancianos institucionalizados con la de aquellos que siguen viviendo en su domicilio o en su entorno familiar.

A) La encuesta general

Ambito y universo

Aunque la *muestra* es prácticamente nacional –pues sólo se excluye a Ceuta y a Melilla–, la población de los municipios más rurales (en cada comunidad autónoma) no forma parte del *universo,* por razones de coste, fundamentalmente. Ahora bien, para

delimitar el margen de mayor ruralidad, teniendo en cuenta la diferente configuración de lo rural y de lo urbano en el territorio español, se considera la densidad y el tamaño medio de los municipios en cada comunidad autónoma, así como la división municipal en entidades singulares de población. Por lo que se decide la siguiente inclusión de municipios por comunidad autónoma (Cuadro 5.14):

CUADRO 5.14. Municipios incluidos en la muestra por comunidad autónoma.

A *Municipios de 1.000 y más habitantes*	*B* *Municipios de 3.000 y más habitantes*	*C* *Municipios de 5.000 y más habitantes*
Aragón Castilla La Mancha Castilla y León Extremadura Navarra La Rioja	Cantabria Cataluña Madrid País Vasco C. Valenciana	Andalucía Asturias Baleares Canarias Galicia Murcia

Estratificación y afijación

Antes de proceder a la selección de los *puntos de muestreo,* se agrupan los municipios mediante el cruce de las 17 comunidades autónomas y el tamaño de la población, dividido conforme a las 9 categorías siguientes:

1) De 1.000 a 3.000 habitantes.
2) De 3.001 a 5.000 habitantes.
3) De 5.001 a 10.000 habitantes.
4) De 10.001 a 20.000 habitantes.
5) De 20.001 a 50.000 habitantes.
6) De 50.001 a 100.000 habitantes.
7) Municipios de más de 100.000 habitantes, no capitales de provincia.
8) Capitales de provincia.
9) Capitales autonómicas.

La *afijación* de las entrevistas a los distintos *estratos* se hace de forma proporcional al número de habitantes *(Censo de población, 1991)* del conjunto de municipios que componen el *estrato*.

Selección de los puntos de muestreo

En cada estrato se eligen los municipios, en función de su tamaño y el número de entrevistas afijadas. El procedimiento es directo, aleatorio y con probabilidad pro-

porcional a la población de cada municipio. Después se extraen, dentro de cada municipio, las entidades singulares de población. Por ello se barajan los criterios siguientes:

a) *Municipios de menos de 10.000 habitantes.* Selección aleatoria de una entidad singular de población, con probabilidad proporcional al número de *habitantes de hecho* registrados en el *Nomenclator de 1986* ó *1981*. Éste es el caso de las provincias de Huelva, Zaragoza, Sta. Cruz de Tenerife, Toledo, Salamanca, Valladolid, Barcelona, Valencia, Orense y La Rioja. La entidad singular capital también entra en el sorteo.

b) *Municipios de más de 10.000 habitantes. Afijación proporcional* de las entrevistas del municipio a la entidad singular capital y al conjunto de entidades singulares no capitales.

Cuando al conjunto de entidades singulares capitales corresponden seis o más entrevistas, se elige aleatoriamente (y con *probabilidad proporcional)* una de ellas, donde se hacen las entrevistas. En la entidad singular capital se realizan las correspondientes o todas, si en el supuesto anterior no se alcanzan las seis entrevistas.

c) *Criterios complementarios*:

1) Considerar entidades singulares elegibles sólo aquellas que cuentan con una población al menos 10 veces superior al número de entrevistas afijadas al municipio.
2) Aunque corresponda aplicar el criterio *b,* por el tamaño del municipio, si éste se compone de "barrios" y "lugares" de peso demográfico parecido al de la entidad singular capital, se utiliza el criterio *a.*

En resumen, han sido 111 los municipios seleccionados (pero 131 las entidades singulares de población visitadas, y por tanto los *puntos de muestreo efectivos);* lo que supone una dispersión aceptable. Más aún, si se tiene en cuenta que la *muestra* toca 42 de las 50 provincias españolas. (Ver cuadro 5. 15.)

Selección de los entrevistados

En el procedimiento de selección de los entrevistados se tienen en cuenta *cuotas* de sexo, edad y condición socioeconómica, tomando como referencia las estimaciones de la *Encuesta de población activa del INE,* correspondientes al tercer trimestre de 1991. Estos son los datos para los grandes subconjuntos de población que han compuesto la *muestra* general (Cuadro 5.16).

CUADRO 5.15. Puntos de muestreo por comunidades autónomas, provincias y municipios, con especificación del estrato al que pertenecen y el número de entrevistas teóricas a realizar.

Comunidad autónoma	*Provincia*	*Municipio*	*Estrato*	*Entrevistas*
Andalucía	Huelva	Aljaraque	3	23
	Cádiz	Algodonales Ubrique San Fernando Algeciras Cádiz	3 4 6 7 8	22 38 39 37 73
	Córdoba	Pozoblanco Córdoba	4 8	37 73
	Granada	Granada	8	73
	Jaén	Martos	5	31
	Málaga	Coín Ronda Estepona Marbella	4 5 5 6	38 40 40 38
	Sevilla	Ginés Montellano Morón Sevilla	3 3 5 9	23 23 40 90
Aragón	Huesca	Biescas Fraga	1 4	11 23
	Zaragoza	Ateca Zaragoza	1 9	12 90
Asturias	Oviedo	Gozón Avilés Oviedo	4 6 8	33 36 60
Baleares	Mallorca	Muro Mahón Palma Mallorca	3 5 9	25 20 40
Canarias	Tenerife	Santiago Teide Candelaria Realejos Laguna Sta. Cruz Ten.	3 4 5 7 8	20 25 49 33 72
Cantabria	Santander	Santoña Santander	4 9	25 30

CUADRO 5.15. *(Continuación)*

Comunidad autónoma	*Provincia*	*Municipio*	*Estrato*	*Entrevistas*
Castilla-La Mancha	Albacete	Balazote	1	14
	Ciudad Real	Retuerta Porzuna Puertollano	1 2 5	13 16 34
	Guadalajara	Guadalajara	8	48
	Toledo	Casarrubios Puebla Almora Sonseca	1 3 3	14 24 25
Castilla y León	Palencia	Palencia	8	51
	León	Boñar Villabalter	1 5	14 35
	Salamanca	Sta. Marta Tor. Salamanca	3 8	19 52
	Valladolid	Fresno Viejo Mojados Tudela Duero Valladolid	1 1 2 9	14 14 16 44
Cataluña	Barcelona	Sta. Coloma C. Argentona Franqueses V. Montcada i R. Rubí Cerdanyola V. Badalona Hospitalet Ll. Barcelona	2 3 4 5 6 6 7 7 9	17 26 35 33 34 35 69 69 216
	Girona	Cassa Selva Salt	3 5	25 34
	Lleida	Lleida	8	39
	Tarragona	Ametlla Mar Deltebre Valls	2 4 5	17 35 34
Comunidad Valenciana	Alicante	Bigastro Muro Alcoy Villena Orihuela Alcoy	2 3 5 5 6	21 24 35 35 46
	Castellón	Castellón Pl.	8	77

CUADRO 5.15. *(Continuación)*

Comunidad autónoma	*Provincia*	*Municipio*	*Estrato*	*Entrevistas*
Comunidad Valenciana	Valencia	Puig	3	24
		Alginet	4	33
		Utiel	4	32
		Untinyent	5	34
		Valencia	9	101
Extremadura	Badajoz	Pucbla Maes.	1	14
		Fuente Cantos	3	17
		Don Benito	5	25
		Badajoz	8	35
	Cáceres	Piornal	1	15
		Talayuela	3	17
Galicia	Coruña	Carreira	3	20
		Arzúa	3	21
		Ribeira	5	29
		Santiago	6	36
	Lugo	Lugo	8	67
	Orense	Celanova	3	20
		Barco Valdeo.	4	33
	Pontevedra	Tui	4	34
Madrid	Madrid	Morata Tajuña	3	27
		S. Fernando H.	5	38
		S. Sebastián R.	6	46
		Fuenlabrada	7	60
		Móstoles	7	62
		Madrid	9	397
Murcia	Murcia	Cehegín	4	34
		Águilas	5	34
		Murcia	9	66
Navarra	Navarra	Doneztebe	1	18
		Barañain	4	17
		Pamplona	9	25
País Vasco	Vizcaya	Muskiz	3	17
		Ermua	4	39
		Galdakao	5	33
		Barakaldo	6	51
	Guipúzcoa	Lazkao	3	17
	Álava	Vitoria	8	97
La Rioja	Logroño	Logroño	9	30

CUADRO 5.16. Estructura sociodemográfica de la muestra.

Características	*Universo * Hab. (miles)*	*Muestra Nºentrev.*	*% Universo y Muestra*
Jóvenes 18-29			
Varones	3.807	623	13,9
Mujeres	3.572	585	13,0
Total	7.379	1.208	26,9
Amas de casa 30-64	4.493	735	16,3
Adultos activos 30-64			
Varones	6.680	1.094	24,3
Mujeres	3.126	512	11,4
Total	9.814	1.605	35,7
Mayores 65 y más			
Varones	2.440	400	8,9
Mujeres	3.363	551	12,2
Total	5.803	951	21,1
Total	27.489	4.500	100,0

* *Fuente: EPA*. Año 1991 (Tercer trimestre).

Todas las entrevistas se realizan, cara a cara, por entrevistadores en los hogares familiares de los núcleos de población elegidos como *puntos de muestreo*. La selección de los hogares se hace conforme a los patrones de zonificación urbanística (centro, ensanche, extrarradio).

Estimación de los errores de muestreo

Debido a la combinación de varios procedimientos muestrales *(estratificación, conglomerados, cuotas)*, no es posible, en rigor, el cálculo de los *errores* derivados del manejo de muestras y del *diseño muestral*. En estas circunstancias, habituales en los sondeos de opinión, se toman como referencia los *errores* que corresponderían a un *muestreo aleatorio simple*. Se trata de una solución sencilla, pero operativa y fundamentada, puesto que respecto al *muestreo aleatorio simple* la *estratificación* mejora el *error* y el *muestreo por conglomerados* lo aumenta.

Por otro lado, conviene recordar que los *errores de muestreo* también dependen de la altura del porcentaje que se analice y del tamaño de las subdivisiones muestrales.

B) Encuesta a mayores en residencias

La *muestra estratégica* a mayores en residencias se circunscribe al territorio de la Comunidad de Madrid. Por ésta y otras razones, los resultados de la investigación se toman como meramente indicativos, incluso para el ámbito madrileño.

La selección de residencias y de entrevistados está guiada por criterios de heterogeneidad (a modo de *muestreo por cuotas*) en las siguientes variables:

a) *Titularidad de las residencias de la tercera edad*. Se tiene en cuenta la distinción entre residencias públicas y privadas, y la distribución porcentual tanto de residencias como de plazas en cada categoría.

   Las cifras de referencia para el año 1988, según la Consejería de Integración Social de la Comunidad de Madrid, cuantifican las residencias privadas en un 89% (con un 58% de las plazas), mientras que las residencias públicas representan los porcentajes complementarios: 11% las residencias, y 42% las plazas.

   En esta *muestra estratégica* son 24 las residencias visitadas. De ellas, 4 son públicas (en las que se realizan el 46% del total de entrevistas) y 20 privadas (donde se hacen el 54% de las 507 entrevistas).

b) *Localización espacial de las residencias*. La diferenciación a este respecto es entre las residencias ubicadas en Madrid capital y las emplazadas en la periferia restante.

   Aunque la fuente citada indica la concentración (65% de las residencias) en Madrid capital, se opta por sobrerepresentar el ámbito periférico a costa del capitalino, en aras de una mayor dispersión espacial de la *muestra*. Además, en términos de plazas, el municipio de Madrid cuenta con menos plazas públicas (32%) que el resto del territorio de la Comunidad de Madrid, según un estudio del Ayuntamiento de Madrid (1988: 98 y ss). Éste ha sido el resultado (Cuadro 5.17):

CUADRO 5.17. Muestra a mayores en residencias, según hábitat.

	Madrid cap.	Periferia	Total
Residencias	8	16	24
Entrevistas	90	387	507

c) *Sexo y edad de los mayores institucionalizados*. La no disponibilidad de datos publicados (con carácter oficial y suficientemente desagregados), hace que se empleen estas dos variables básicas de forma aproximada. Para ser precisos, no se considera como población residente para entrevistar a los mayores deno-

minados "no válidos" o "asistidos" (aquéllos necesitados de mayores cuidados), debido a la dificultad mayor para entrevistarles. Sin embargo, se decide dar *cuotas* de sexo y edad a los entrevistadores, con la finalidad de evitar el *sesgo muestral* hacia los residentes más jóvenes, con mejor salud o más dispuestos.

En el Cuadro 5.18 se describe la *muestra a mayores en residencias*.

CUADRO 5.18. Mayores entrevistados en residencias, según edad y sexo.

Edad	*Ambos sexos*		*Varones*		*Mujeres*	
–78 años	158	31%	64	36%	94	29%
78 y +	349	69%	114	64%	235	71%
Total	507	100%	178	100%	32	100%

Lecturas complementarias

Azorín, F. y Sánchez Crespo, J. L. (1986): *Métodos y aplicaciones del muestreo*. Madrid, Alianza.

Henry, G. T. (1990): *Practical sampling*, Newbury Park, Sage.

Levy, P. y Lemeshow, P. (1991): *Sampling of population: methods and applications*. New York, Wiley.

Moser, C. A. y Kalton, G. (1989): *Survey methods in social investigation*. Aldershot, Gower Publishing Company.

Rodríguez Osuna, J. (1991): *Métodos de muestreo*, Cuaderno Metodológico n.° 1. Madrid, CIS.

Rodríguez Osuna, J. (1993): *Métodos de muestreo. Casos prácticos*, Cuaderno Metodológico n.° 6. Madrid, CIS.

Ejercicios Propuestos

1. Compare los dos diseños muestrales desarrollados en el apartado 5.5, destacando las diferencias y semejanzas entre ellos.
2. Se desea conocer la opinión de los jueces sobre el nuevo Código Penal. Concrete las fases del diseño de la muestra.
3. Un equipo de investigadores desea entrevistar a enfermos de sida para analizar la repercusión de la enfermedad en su vida diaria. ¿Qué diseño muestral se adecuaría más a este objetivo de investigación? Especifique cómo seleccionaría las unidades muestrales.
4. En un municipio de 185.000 habitantes se extrae una muestra de 10.000 individuos para conocer la renta mensual por habitante en dicho municipio. Ésta resulta ser de 95.000 ptas al mes, con una desviación típica de 3.600. Calcular el error que representa la muestra y los intervalos de confianza. Afijar la muestra e indicar qué tamaño se precisaría para un error máximo del ±3%.

Clase social	*Porcentaje*	*Desviación*
Alta	37,7	2.500
Media	39,9	2.100
Baja	22,4	1.600

5. A continuación se exponen los datos correspondientes a una encuesta de 2.000 entrevistas a jóvenes entre 18 y 30 años. ¿Qué tipo de afijación se ha efectuado? Pondere la muestra, si lo estima necesario.

Hábitat	*Universo % (por mil)*	*Muestra*	*A favor de ampliar el aborto*
Urbano	404	1000	75
Intermedio	364	600	67
Rural	232	400	58

TERCERA PARTE

LA OBTENCIÓN DE INFORMACIÓN

6

EL USO DE FUENTES DOCUMENTALES Y ESTADÍSTICAS

Una *estrategia* básica de investigación consiste en acudir a archivos de datos y a fuentes bibliográficas en busca de la información, que otros autores han reunido, concerniente al problema de estudio. Esta *estrategia* se encuadra dentro de la denominada "*investigación no-reactiva*", a partir de la publicación de Webb *et al.* (1966), *Unobtrusive Measures: Nonreactive Research in the Social Sciences*.

Esta obra presenta una serie de enfoques de investigación dirigidos a minimizar los efectos, que la presencia del investigador puede ocasionar, en la obtención de información. Comprende los *archivos documentales y estadísticos*, además de los *documentos personales* (autobiografías, diarios, memorias, cartas, fotos o cualquier registro icónico u objeto personal) y la *observación no interactiva*. De ellos, en este capítulo se tratarán, exclusivamente, el *uso de fuentes (archivos) documentales y estadísticos*. Para una introducción al *uso de la documentación* y la *observación-participación* en la *investigación cualitativa*, remito al lector interesado a la lectura de los capítulos 4 y 5 del manual anteriormente referido de Miguel Valles en esta misma colección.

La exposición de esta *estrategia de investigación* antecede –en este manual– al resto de las *estrategias* aquí descritas *(encuesta y experimentación)*, por su consideración de "básica". Esto es, presente en la mayoría de los estudios, bien como *estrategia* única de investigación, o bien complementando otras *estrategias* principales diseñadas para alcanzar los objetivos del estudio.

Como ya se ha razonado en la enumeración y acotación de las *estrategias de investigación* (Capítulo 3), la *revisión bibliográfica* de investigaciones teóricas y empíricas sobre el tema concreto de estudio (y otros similares) figura en los preliminares de cualquier indagación. Fundamentalmente, porque contribuye a:

a) La *familiarización* con el tema de estudio, sus antecedentes y la metodología ya ensayada.

b) La *estructuración* de las ideas originarias del estudio en un *diseño de investigación* concreto, al indicar:

1) Aspectos a tratar *(hipótesis* a comprobar).
2) Sujetos a analizar (características de la *población* de interés).
3) *Estrategias* y *técnicas* de recogida y de análisis a aplicar, tras los resultados y experiencia adquirida en indagaciones precedentes.

La información así obtenida (mediante *fuentes secundarias)* descubrirá el conocimiento que ya se tiene sobre el tema elegido frente a los aspectos necesitados de mayor profundización a través de investigación "primaria". Como afirma Stewart (1984: 13):

> "El propósito de la investigación primaria generalmente debería ser rellenar los agujeros en el conocimiento existente. Estos agujeros no pueden ser identificados sin una comprensión de la base de conocimiento existente. Es quizás desafortunado que el término "secundaria" se haya elegido para referirse a datos existentes. Este término no implica algo sobre la importancia de la información."

Pero, ¿qué se entiende por *investigación secundaria*?, ¿qué términos comprende?

6.1. Clarificación terminológica: el análisis secundario y el meta-análisis

En general, por *investigación primaria* se entiende cualquier tipo de indagación en la que el investigador analiza la información que él mismo obtiene, mediante la aplicación de una o varias *técnicas* de obtención de datos *(cuestionario, guión de entrevista, observación)*. Por el contrario, la *investigación secundaria* se limita al análisis de datos recabados por otros investigadores (con anterioridad al momento de la investigación). Sin embargo, ambos tipos de investigación no constituyen modalidades contrapuestas, sino complementarias. La *investigación secundaria* se considera extensión y punto de partida habitual de la *indagación primaria.*

Dentro de la *investigación secundaria* Hakim (1994) diferencia tres variedades importantes:

a) La revisión de investigaciones.
b) El meta-análisis.
c) El análisis secundario.

• *La revisión de investigaciones*

Representa uno de los preliminares esenciales en cualquier indagación empírica, al proporcionar una *síntesis* del conocimiento existente sobre un tema específico.

Esta "síntesis" resulta de la valoración de la información disponible y su adecuación a los propósitos de la investigación.

La *revisión* varía en *énfasis* (teórico y/o metodológico) y en el período de *tiempo* que comprenda: incluye sólo estudios coetáneos o puede abarcar diferentes momentos temporales, proporcionando así una perspectiva histórica a la investigación.

El número de estudios revisados también varía dependiendo del *tema* que se analice y de la *experiencia* del investigador en la localización del material relevante (Hakim, 1994).

• *El meta-análisis*

Se define como "el análisis de los análisis (el análisis estadístico de los hallazgos de muchos análisis individuales)" (Glass *et al.,* 1987: 21). "No es un método de investigación o una fuente de datos", sino que "proporciona un medio de resumir los resultados de numerosos estudios cuantitativos sobre un dominio particular" (Bryman, 1995: 228).

Tiene como objetivo principal la obtención de un cuadro resumen, que sintetice los resultados cuantitativos alcanzados en distintas investigaciones. Cada estudio revisado constituye una unidad de la *muestra de estudios* (relevantes), extraída del total existente. En ella se aplican distintas *técnicas de análisis estadístico (univariable, bivariable o multivariable)*, para comprobar el *nivel de significatividad estadística* de determinados hallazgos.

La consecución de este objetivo supone:

a) La *descripción* de los hallazgos y cómo varían de un estudio a otro. Los hallazgos se convierten en la *variable dependiente* (en los *análisis estadísticos),* mientras que las características (sustantivas y metodológicas) de los estudios en las *variables independientes.*
b) La comprobación de su *significatividad.*
c) La determinación de su *magnitud* conjunta.

Esta *cuantificación* suele implicar, de acuerdo con Glass *et al.* (1987: 69):

a) "Medición en sus aspectos métricos (por ejemplo, ¿en qué año se realizó el estudio?, ¿cuál es el tamaño de la *muestra* en la que se basa el *coeficiente de correlación* r_{xy} obtenido?).
b) Su función nominal o de codificación (por ejemplo, ¿fueron las diferencias iniciales corregidas por el *análisis de la covarianza*? sí = 1, no = 2)".

A ello se suma la aplicación de *tests de significatividad estadística,* como la *t de Student* o la X^2.

La solidez de las conclusiones del *meta-análisis* depende de la calidad de los estudios revisados, de si "representan" realmente a la totalidad de estudios existentes (si se incluyen tanto investigaciones publicadas como no publicadas), y de cómo se haya procedido en el análisis de cada uno de ellos.

• *El análisis secundario*

"Es sencillamente un análisis posterior de la información que ya se ha obtenido. Tal análisis puede estar relacionado con el propósito original para el que los datos se recogieron, o puede dirigirse a un asunto bastante diferente del que instó el esfuerzo de reunión de los datos originales. Puede implicar la integración de distintas fuentes o un reanálisis de los datos de una fuente única" (Stewart, 1984: 11).

El *análisis secundario* facilita el *análisis comparativo* y el *de tendencias,* a partir de los datos disponibles para un amplio período de tiempo. De acuerdo con Hakim (1994: 23):

> "El análisis secundario de datos existentes es probable que siga siendo la aproximación más corriente para realizar estudios comparativos internacionales, especialmente para estudios que buscan cubrir un número elevado de países y/o tendencias a lo largo del tiempo."

De hecho, una proporción importante de "investigación económica se basa en el análisis secundario de series temporales de nivel macro, que consisten en un gran número de indicadores estadísticos nacionales y medidas cotejadas desde una gran variedad de encuestas oficiales y series estadísticas" (Hakim, 1994: 22).

Sin duda, estos tipos de análisis se han visto favorecidos por la eclosión, en los últimos años, de las aplicaciones informáticas para la detección y utilización de la *información secundaria.*

6.2. Fuentes de información "secundaria"

La información *secundaria* engloba tanto datos "brutos", elaborados por distintos organismos (públicos o privados) para sus propios propósitos, como los proporcionados y analizados en distintas publicaciones. El amplio abanico existente puede resumirse en los apartados siguientes:

1) *Datos no publicados,* elaborados por organismos públicos y privados, relativos a su actuación.
2) *Datos publicados* por organismos públicos y privados: estadísticas e informes.
3) *Investigaciones publicadas* en libros y revistas.
4) *Investigaciones no publicadas.*

6.2.1. *Datos no publicados, elaborados por organismos públicos y privados, relativos a su actuación*

Todo organismo (o empresa) registra alguna información sobre su personal y aspectos relacionados con su actividad. Esta información adquiere un gran interés en la descripción y evaluación de la actuación de estos organismos. Pueden citarse, a modo de ejemplo, los *expedientes* incoados por instancias judiciales, los *informes* de pacientes de centros sanitarios, los *expedientes académicos* de alumnos de centros de enseñanza, los *registros* de afiliados a asociaciones o del personal de una empresa.

El carácter privado (no "público") de este tipo de información obstaculiza su acceso. En caso de requerirla, el investigador tendrá que solicitarla, expresa y formalmente, al organismo en cuestión. En esta solicitud se ha de enfatizar la trascendencia de la información requerida para los objetivos de la investigación; además de garantizar el anonimato de las personas a las que se refieran los datos manejados.

6.2.2. *Datos publicados por organismos públicos y privados: estadísticas e informes*

Algunas entidades (públicas y privadas) editan *estadísticas* e *informes* como parte principal (o complementaria) de su actividad. La *validez* y *fiabilidad* de estas fuentes es algo que el investigador deberá comprobar contrastándola, por ejemplo, con la información que recabe de dichos organismos mediante otras fuentes: bien vía *investigación primaria* (una *encuesta,* por ejemplo) o *secundaria (estadísticas y/o informes* elaborados por otras instancias). La razón principal de esta necesaria contrastación reside en la posibilidad de que el organismo productor distorsione sus *informes* para crear una buena imagen pública de su actuación.

EJEMPLO DE CONTRASTACIÓN DE DATOS PUBLICADOS

En la *evaluación* de la intervención de los Tribunales Tutelares de Menores (Cea, 1992), ya relatada –véase Capítulo 3–, se hizo uso de distintas *fuentes documentales y estadísticas.* Concretamente, se contrastaron *Estadísticas del Tribunal Tutelar de Menores de Madrid* (referentes a los ejercicios de 1975 hasta 1983), con los datos publicados en las *Estadísticas judiciales* del Instituto Nacional de Estadística (del mismo período de tiempo). Estos *datos secundarios* se contrastaron, a su vez, con la información registrada en los *expedientes* de menores incoados en las susodichas fechas.

La información reunida de las tres *fuentes* citadas no coincidía, ni siquiera las cifras estadísticas proporcionadas por el Tribunal de Menores de Madrid y las publicadas por el INE (pese a que este último organismo confecciona sus *Estadísticas judiciales* de los boletines que periódicamente le remiten los tribunales de justicia).

Entre las causas de esta falta de concordancia en la información procedente de distintas fuentes, se observaron *errores en el registro* de la información (desde el mismo momento en

que se abre expediente al menor), junto a la existencia de *criterios de clasificación* desiguales en ambos organismos. A ello se suman las incorrecciones encontradas en la actuación del personal de los tribunales en el cumplimiento de su función.

El conocimiento del proceso seguido en la elaboración de las *estadísticas* ayudó en su interpretación.

Además de las *estadísticas e informes publicados* por entidades particulares, por parte de las administraciones públicas se realizan –con mayor o menor rigor– "estadísticas oficiales". Estas describen, con cierta regularidad temporal, algunas de las características sociodemográficas de la población.

En España, a *nivel nacional,* son de obligada referencia las estadísticas elaboradas y publicadas por el Instituto Nacional de Estadística (INE). De ellas cabe mencionar las siguientes:

- Censo de población (desde 1857).
- Anuario estadístico de España (desde 1858).
- Nomenclátor (desde 1858).
- Movimiento natural de la población (desde 1858).
- Censo de edificios y viviendas (desde 1950).
- Estadísticas judiciales (desde 1959).
- Proyección de la población española (desde 1978).
- Encuesta de migraciones (desde 1980).
- Migraciones (desde 1985).
- Encuesta Continua de presupuestos familiares (desde 1985).
- Censo electoral. Informe sobre la revisión anual (desde 1992).
- Clasificación nacional de actividades económicas (desde 1993).
- Boletín mensual de estadística.
- Indicadores sociales.
- Panorámica social de España.
- Encuesta sociodemográfica 1991.
- Encuesta de fecundidad.
- Encuesta de salarios en la industria y los servicios.
- Estudio sobre la nutrición (basado en la Encuesta de presupuestos familiares 1990-1991).

Además de estas publicaciones impresas (algunas de ellas también disponibles en soporte magnético), el INE oferta otros servicios de difusión de estadísticas:

a) El *banco de datos TEMPUS,* de acceso gratuito a través de la red telefónica conmutada o la red IBERPAC (mediante un módem conectado a un PAC de X.28 o a un nudo de X.25).

Este *banco de datos* reúne una colección de más de 400.000 series cronológicas (mensuales, trimestrales y anuales) de diversas fuentes: Instituto Nacional de Estadística, Ministerio de Trabajo y Seguridad Social, Banco de España, Ministerio de Industria y Energía, principalmente. A ellas se suman diversas series internacionales de la colección *Main Economic Indicators* de la OCDE.

b) El *banco de datos VIDEOTEX*, al que se accede marcando 031 en la red IBERTEX, y tecleando *INE#. Con este servicio se dispone de los datos del INE al momento y a un bajo coste (la llamada telefónica).

El Banco de España y los distintos ministerios también presentan un amplio catálogo de publicaciones, sobre materias afines a su función. Tal es el caso de las estadísticas editadas por el Ministerio de Trabajo y Seguridad Social:

- Anuario de estadísticas laborales.
- Estadísticas de accidentes de trabajo y enfermedades laborales.
- Estadísticas de huelgas y cierres patronales.
- Estadísticas de permisos de trabajo a extranjeros.

Asimismo, han de mencionarse los *Boletines Estadísticos de Datos Básicos* del Ministerio de Asuntos Sociales, presentado en papel y en disquete (para facilitar su tratamiento informático).

A *escala regional,* desde los años ochenta, algunas *Comunidades autónomas* han creado institutos de estadística propios (País Vasco, Andalucía, Valencia, por ejemplo). En otros casos (como Madrid), la producción regional de información estadística (demográfica y económica, sobre todo) se ha organizado en un departamento de estadística, normalmente dependiente de la Consejería de Economía.

Entre las estadísticas demográficas publicadas por la Consejería de Economía de la Comunidad de Madrid, por ejemplo, están las siguientes:

- Anuario estadístico de la Comunidad de Madrid (desde 1984).
- Censo de población de los municipios de menos de 50.000 habitantes de la Comunidad de Madrid.
- Censo de viviendas de los municipios de menos de 50.000 habitantes de la Comunidad de Madrid.
- Resultados adicionales del censo de población de la Comunidad de Madrid.
- Proyecciones de población de la Comunidad de Madrid a 1996.
- 2001: Proyecciones de población de la Comunidad de Madrid.
- Proyección de población y hogares de la Comunidad de Madrid con horizonte en el año 2006.
- Encuesta demográfica de la Comunidad de Madrid (de 1986 y de 1991).

- Padrón de habitantes de 1986 de la Comunidad de Madrid.
- Nomenclátor de la Comunidad de Madrid 1991.
- Estadísticas del movimiento natural de la población de la Comunidad de Madrid (desde 1986), en colaboración con el Ayuntamiento de Madrid y la Consejería de Salud.
- Boletín trimestral "demografía y salud" (desde enero de 1988).
- Censos de población y vivienda de 1991 de la Comunidad de Madrid.
- Clases sociales: estudio comparativo entre España y la Comunidad de Madrid 1991.
- Actitudes y representaciones sociales de la población de la Comunidad de Madrid en relación con los censos de población y vivienda de 1991.
- 65 años cumplidos: los ancianos de la Comunidad de Madrid.

A estas publicaciones se suman los *Documentos de trabajo* (también editados por el Departamento de Estadística de la Comunidad de Madrid) y los *Ficheros Tipo Matriz, por ámbitos geográficos* (datos en soporte magnético) siguientes: los censos de población y vivienda de 1991 de la Comunidad de Madrid; las Estadísticas de movimiento natural de la población de la Comunidad de Madrid; y los Resultados electorales de la Comunidad de Madrid.

A *escala municipal,* algunos ayuntamientos (como el de Madrid y Barcelona, entre otros) también generan información estadística desagregada para su ámbito municipal (por distritos y barrios administrativos, por ejemplo). No se trata solamente de explotaciones detalladas del *Padrón municipal de habitantes,* sino también de estudios sociológicos hechos por el personal técnico del Ayuntamiento, o encargados a profesionales externos. Entre estos estudios cabe citar la *Encuesta sobre formas de vida del Ayuntamiento de Madrid.*

Respecto a *entidades privadas*, han de citarse las publicaciones de los servicios de estudios de distintas entidades bancarias, como BANESTO, el Banco Bilbao-Vizcaya o la Confederación de Cajas de Ahorros, por ejemplo.

También es obligado hacer referencia a los estudios promovidos por fundaciones, como la Fundación FOESSA (Fomento de Estudios Sociales y de Sociología Aplicada), que desde 1969 viene patrocinando distintos informes sociológicos sobre la situación social en España. De más reciente creación es el Centro de Investigación de la Realidad Social (CIRES), que desde 1990 lleva a cabo diferentes encuestas sobre aspectos sociológicos.

A *escala internacional* destacan (en el contexto europeo) las "estadísticas oficiales" efectuadas por EUROSTAT (Oficina Estadística de las Comunidades Europeas). Entre sus publicaciones estadísticas se encuentran las siguientes:

- Anuario Eurostat'95. Visión estadística sobre Europa 1983-1993.
- Estadísticas básicas de la Unión Europea.

- Informe anual 1994 (Comité Económico y Social).
- *Eurostat Yearbook 95. A Statistical Eye on Europe 1983-1993.*
- *Statistiques démographiques 1995.*
- Estudios de desarrollo regional –Estudio prospectivo de las regiones del Mediterráneo Oeste. Europa 2000–.
- Europa 2000+. Cooperación para la ordenación del territorio europeo.

Este mismo organismo cuenta, además, con varias *bases de datos*. De ellas son de especial referencia:

- EUROSTAT-CD (que contiene estadísticas sociales y económicas, datos regionales y sobre comercio exterior).
- CD-ROM COMEXT (base de datos comunitarias sobre el comercio exterior).
- CD-ROM Panorama de la industria comunitaria (proporciona una amplia panorámica de la situación de la industria y los servicios de la Unión Europea).
- CD-ROM REGIOMAP (que integra las estadísticas regionales oficiales de EUROSTAT procedentes de la base de datos REGIO –Estadísticas regionales– con los datos topográficos y los límites territoriales de GISCO en un *software* de estadísticas y cartografías).

También cabe mencionar aquí los informes de la Organización de Cooperación y Desarrollo Económico (OCDE) sobre distintos temas, y sus *bases de datos* como, por ejemplo, *CD-ROM OECD Statistical Compendium* (que proporciona más de 150.000 series cronológicas, actualizadas anualmente, seleccionadas de las principales bases de datos estadísticos de la OCDE).

De Naciones Unidas son conocidos el *Statistical Yearbook* (también en CD-ROM) y el *Demographic Yearbook,* entre otras publicaciones. Su base de datos *STATBASE LOCATOR ON DISK* (disponible en disquetes) facilita la localización de bases de datos estadísticas internacionales en soporte electrónico.

La UNESCO publica igualmente un Anuario estadístico. Referencias a documentos y publicaciones de la UNESCO pueden encontrarse en CD-ROM Bases de Datos de la UNESCO.

En este breve repaso de fuentes principales de datos estadísticos, merecen asimismo resaltarse los Informes sobre el Desarrollo Mundial, del Banco Mundial. Este organismo también edita soportes magnéticos sobre investigación, análisis e información estadística de distintos aspectos del desarrollo económico y social. Entre sus publicaciones electrónicas están:

- *World Tables 1994. Disquetes* (de datos económicos, demográficos y sociales sobre más de 160 economías, así como indicadores básicos de otras 50).

- *Social Indicators of Development 1994. Disquetes* (que incluye estimaciones sobre fertilidad, mortalidad, analfabetismo y salud, entre otras).
- *CD-ROM World Data* (el primer CD-ROM elaborado por el Banco Mundial, que contiene datos estadísticos relativos al período 1960-1992).
- *CD-ROM World Development Report 1978-1995* (que constituye un archivo de todos los *World Development Reports* publicados por el Banco Mundial desde 1978 hasta 1995).

6.2.3. Investigaciones publicadas

Existen varias formas de acceder a estudios publicados. Tradicionalmente, la manera más rápida era recurrir a *fuentes secundarias* que citan y resumen investigaciones publicadas en distintas áreas de conocimiento. Las principales *fuentes de referencias* son:

a) *Índices* como *The Social Sciences Citation Index* (SSCI), que compila unos 130.000 artículos de revista cada año. A su vez contiene un índice temático: el *Permuterm Subject Index.*

Otros índices que cubren bibliografía especializada son: *Current Index to Journals in Education, Index of Economic Articles, Social Science Index, Population Index* o *The International Population Census Bibliography: revision and updata,* entre otros.

b) *Abstracts*, como por ejemplo: *Sociological Abstracts, Psychological Abstracts, Abstracts in Antropology, Dissertation Abstracts International, International Political Science Abstracts,* por ejemplo.

Estos *abstracts* contienen breves resúmenes (en torno a 100 palabras redactadas por el propio autor del estudio) de investigaciones publicadas en distintos países del mundo, de las que se tiene conocimiento por la revista que los edita. Las investigaciones figuran ordenadas por tema y autor.

Sin embargo, la generalidad de los *index* y *abstracts* presentan –siguiendo a Cooper (1984)– tres *límites* importantes:

a) La larga carencia temporal (con frecuencia más de dos años) desde que el estudio concluye y éste aparece registrado en el *abstracts*.

b) Se centran en disciplinas particulares, cuando las investigaciones son, con frecuencia, interdisciplinales.

c) Su organización. Aunque el *abstract* cubra, con exhaustividad, las revistas relevantes a un tema concreto, es probable que algunos artículos se pierdan, por no ser fácilmente localizados. Esto último ocurre cuando el investigador no conoce términos clave de los *índices* que se aplican a artículos, o cuando el autor de los *índices* omite términos que el investigador emplea comúnmente.

De ahí que no sorprenda que algunos autores (Glass *et al.*, 1987) afirmen que es difícil conocer si los estudios que figuran en los *index* o *abstracts* representan la mayoría de la evidencia existente sobre la cuestión que se investiga, o sólo una parte no representativa de la misma.

Junto con los *index* y *abstracts*, están los *Reviews* (como *Sociological Review, Review of Educational Research*, o *Psychological Bulletin*, por ejemplo). Estos comprenden artículos y recensiones de publicaciones (libros y revistas) sobre un tema particular. Recensiones de publicaciones también pueden encontrarse en la generalidad de las revistas científicas editadas.

Una vez localizados los estudios (artículos o libros) de interés, podrán conocerse otras *referencias bibliográficas* (no proporcionadas en los *abstracts*), gracias a las bibliografías adjuntas en las publicaciones consultadas. A su vez, éstas remitirán a nuevas referencias y así, sucesivamente, hasta que el investigador decida dar por concluida la *revisión bibliográfica*.

Los grandes avances en el campo de la informática han propiciado la consecución de una *búsqueda bibliográfica* más rápida y cómoda (vía terminales de ordenador o microordenadores), que los métodos tradicionales anteriormente referidos. La mayoría de las grandes bibliotecas disponen hoy de terminales de ordenador, que permiten acceder tanto a *bases de datos* externas, como a la utilización de *bases de datos* en discos compactos tipo CD-ROM.

Cada disco CD-ROM contiene una *base de datos*, o una parte de ella (como PSYCLIT, por ejemplo, que comprende varios volúmenes del *Psychological Abstracts*). La editorial Sage (una de las editoriales internacionales más importantes en la divulgación de literatura metodológica) ofrece también una *base de datos de Metodología de la Investigación Social* en CD-ROM. Esta base permite la búsqueda de más de 34.000 *referencias bibliográficas* claves en la metodología de las ciencias sociales desde un PC, en una fracción de tiempo considerablemente inferior al necesario para una búsqueda convencional en bibliotecas.

Pero, como la utilización de los CD-ROM también está condicionada al empleo de palabras "clave" (como en los *abstracts* e *index*), su uso no logra solventar las deficiencias observadas en la búsqueda bibliográfica limitada a la consulta manual de los *abstracts*. Estas *deficiencias* se resumen en dos fundamentales:

a) Que las *palabras clave* no identifiquen *referencias bibliográficas* de interés.
b) Que la *base de datos en CD-ROM* se restrinja a *referencias bibliográficas* tomadas de unas fuentes específicas, y en fechas relativamente recientes. Así. Stewardt (1984) observa que la mayoría de las *bases de datos* no contienen *referencias bibliográficas* previas a 1970.

6.2.4. *Investigaciones no publicadas*

La fuente principal de los datos no publicados son los *archivos o bancos de datos secundarios* (de encuesta, generalmente). La mayoría de estos *bancos de datos*

pertenecen a universidades y centros de investigación. Entre los más conocidos están:

- El *Inter-University Consortium for Political and Social Research* (ICPSR), en la Universidad de Míchigan (EEUU).
- *El Economic and Social Research Council* (ESRC), en la Universidad de Essex (Inglaterra).
- Los *Steinwetz Archives*, en Holanda.

En España, el Instituto de Información y Documentación en Ciencias Sociales y Humanidades (ISOC), dependiente del CSIC (el Consejo Superior de Investigaciones Científicas) dispone de un servicio especializado en información bibliográfica. Este servicio comprende la edición de repertorios bibliográficos y la creación de *bases de datos* automatizadas.

El Centro de Investigaciones Sociológicas (CIS), dependiente del Ministerio de la Presidencia, también comprende un *banco de datos* que reúne información de las numerosas encuestas llevadas a cabo en dicho centro desde su fundación.

En la vertiente privada, hay nuevamente que hacer referencia al CIRES, si bien este centro de investigación ofrece un volumen de información considerablemente inferior, debido a su reciente creación.

En la actualidad, la red INTERNET representa la vía más rápida y económica de acceder a información disponible en bibliotecas y bancos de datos. Además permite solicitar la remisión de los estudios que se precisen.

Por último, los congresos, seminarios, o *simposia,* permiten también el conocimiento y divulgación de investigaciones no publicadas, realizadas en fechas recientes. Estas investigaciones figuran reseñadas en los *resúmenes de comunicaciones y ponencias* editados por el organismo responsable del congreso o simposia. Véase, por ejemplo, los publicados por la Federación Española de Sociología de los últimos congresos celebrados.

6.3. Evaluación y análisis de datos secundarios

Antes de proceder al análisis de la información adquirida mediante *fuentes secundarias,* hay que evaluar la calidad de la misma. En conformidad con Stewart (1984), la evaluación de *datos secundarios* debería seguir los mismos procedimientos aplicados en la evaluación de *datos primarios*. Éstos se resumen en los aspectos siguientes:

a) *Fuentes de datos:* ¿cuál era el propósito del estudio?, ¿quién recogió la información?

b) *Medidas utilizadas:* la calidad de los *datos secundarios* no puede evaluarse sin un conocimiento previo de la metodología empleada cuando se recopilaron los datos. Específicamente, se precisa información sobre:

1) El *diseño muestral,* con especial atención a la representatividad de la *muestra*.
2) Los *porcentajes de respuesta* conseguidos.
3) Las *técnicas de obtención de datos* empleadas (de observación, de entrevista, de documentación).
4) Las *técnicas analíticas* aplicadas a la información reunida.

c) *El tiempo de recogida de los datos*. Cuando se utilizan *datos secundarios* hay que cerciorarse de la fecha de obtención de la información.

Por *ejemplo,* la referencia temporal de los datos de una *encuesta,* en la que se basa un informe sociológico, dará la fecha de realización del *trabajo de campo*. Otro ejemplo se tiene en los *censos de población,* cuyas estadísticas van referidas a una fecha determinada: antes a 31 de diciembre del año censal, ahora a 1 de marzo.

Cualquier investigador sabe del desfase entre la recogida de los datos y la publicación de los resultados. Este desfase, sumado al tiempo transcurrido hasta el momento en el que se baraja la utilización de los datos, puede hacernos perder el interés por éstos. No quiere ello decir que los datos de hace unos años hayan perdido todo su valor para el sociólogo. Ésa sería una concepción errónea de la división del trabajo entre sociólogos e historiadores. Piénsese en lo imprescindibles que se vuelven los datos que se obtuvieron tiempo atrás para analizar los más recientes con la necesaria perspectiva histórica.

d) *La adecuación de los análisis y las conclusiones*. Habrá que comprobar la consistencia de la información con la proporcionada por otras fuentes.

Cuando la misma información se presenta mediante varias fuentes de datos independientes, la confianza en esos datos aumenta. De ahí que se recomiende el contraste de información, al menos de dos fuentes distintas.

En caso de no coincidencia en la información, habría que identificar posibles razones de la discrepancia observada y determinar qué fuente se estima más fiable.

MacDonald y Tipton (1993) agrupan las cuestiones de evaluación e interpretación en cuatro criterios esenciales en todo *análisis documental:*

a) *Autenticidad:* si el documento es o no auténtico. En caso de serlo, si lo es en su totalidad o sólo en una parte.
b) *Credibilidad* de los datos y de la fuente consultada.
c) *Representatividad* respecto del conjunto de documentos existentes.

d) *Significado* aparente y profundo. Para esto último se emplearían, preferentemente, técnicas analíticas cualitativas, como el *análisis estructural* o el *interpretacional*.

Concluida la evaluación, comienza el *análisis de la información*. Éste puede consistir sólo en la *lectura analítica* (cuando se realizan revisiones de investigaciones), o bien en la aplicación de *análisis de contenido* (cuantitativo y/o cualitativo) en el análisis de documentos, o ir más allá, empleando técnicas estadísticas complejas (especialmente el *meta-análisis)*.

En general, pueden emplearse las mismas *técnicas analíticas* utilizadas para la descripción e interpretación de la *información primaria*. Desde *análisis exploratorios* (de distribuciones de frecuencias y descripción univariable) hasta *análisis multivariables* sofisticados. En los Capítulos 9 y 10 se describe la diversidad analítica existente en la *investigación cuantitativa (primaria y secundaria)*. Para su conocimiento, remito a la lectura de dichos capítulos.

6.4. Ventajas e inconvenientes del uso de fuentes documentales y estadísticas

La principal *ventaja* de la *información secundaria* es su disponibilidad y a bajo coste, especialmente si la información se halla publicada. En caso contrario, el acceso a dicha información resultará más arduo.

El recurso a esta estrategia de investigación se recomienda cuando:

a) Se desea información sobre acontecimientos del pasado, y ésta es difícil de obtener por otros medios (por no localización de sus actores, o porque éstos no la recuerden con exactitud).
b) Se diseña una *investigación primaria*. La *revisión bibliográfica* ayuda a la formulación del *problema,* planteamiento de las *hipótesis,* acotación de la *población* a analizar, y la selección de la *metodología* a emplear (a partir de la experiencia alcanzada en investigaciones anteriores).
c) Se dispone de recursos (económicos, materiales y humanos) limitados para desarrollar una *investigación primaria (trabajo de campo)*.
d) Se precisa dar una visión general de un fenómeno social concreto, incluyendo distintos entornos socio-culturales y períodos de tiempo.

El recurso a *información secundaria* se muestra de gran utilidad en *estudios comparativos* (nacionales e internacionales) y *de tendencias*. Si bien, ello exige que se hayan seguido unos mismos criterios para la obtención de información en las fuentes originales (que se haya, por ejemplo, aplicado un mismo *cuestionario* en distintos países o períodos de tiempo). En caso contrario, no podrán compararse los resultados obtenidos.

A este respecto Stewart (1984: 113) expone que "no es inusual que se tomen prestados cuestionarios e items de los ya existentes en la literatura. Esto no sólo reduce el trabajo requerido para desarrollar un nuevo instrumento de investigación, sino que también permite mayor comparabilidad entre la investigación previa y el nuevo estudio. Es costumbre citar la fuente de la que se toman los *ítems*. Cuando se toman los instrumentos enteros, normalmente se debe obtener permiso del autor".

e) En el *diseño muestral,* al proporcionar los *datos censales* características de la *población* de interés. Esta información resulta necesaria para adecuar la *muestra* a la *población* y, de esta forma, alcanzar una mayor representatividad y precisión de las *estimaciones muestrales*.

f) En los *diseños cuasi-experimentales,* cuando se comprueban los efectos de intervenciones o *tratamientos concretos,* antes y después de su aplicación.

g) En cualquier *investigación multimétodo,* como complemento a la información recabada mediante otras *estrategias de investigación*.

En su contra, el principal *inconveniente* del *uso de información secundaria disponible* es su no adecuación. Ello puede deberse bien a la inexistencia de la información que se desea (sobre todo, si ésta es muy específica), o bien a que la información existente no se adecúa a los objetivos concretos de la investigación. Los *datos secundarios* pueden ser incompletos o haber sido recogidos para otros propósitos diferentes de los perseguidos por el investigador que los consulta.

A la no adecuación se suman otros *inconvenientes* fundamentales:

a) Como generalmente transcurre un largo período de tiempo desde que se recoge la información original y ésta es analizada y publicada, suele haber una demora considerable en la disponibilidad de los *datos secundarios*.

La demora llega, por ejemplo, a ser de tres años en las *Estadísticas judiciales* elaboradas por el INE: en 1996 se conoce la información relativa al ejercicio de 1993.

De ahí que se considere a los *datos secundarios* "por definición, viejos" (Stewart, 1984: 14). Lo que repercute, lógicamente, en su no conveniencia para la descripción de fenómenos contemporáneos.

b) Cuando se utilizan *datos secundarios* no siempre se tiene información que ayude a evaluar la calidad de los datos proporcionados. Concretamente, el conocimiento de los *errores muestrales,* los *errores de medición* (operacionalización de conceptos e interpretación), y los de *obtención y análisis* de la información.

Su conocimiento sin duda ayudará no sólo a la evaluación de la calidad de la información registrada, sino también a su interpretación. De ahí la importancia de obtener una copia del instrumento de medición empleado, junto con las especificaciones técnicas de su aplicación.

c) Las *fuentes secundarias* disponibles pueden reunir información únicamente de algunas de las variables que el investigador precise conocer. También puede acontecer que estas variables hayan sido medidas de forma no coincidente con los intereses del investigador.

Los *datos estadísticos* se presentan, habitualmente, agrupados en un número concreto de categorías, y para una población determinada. Si el investigador desease información individual o desagregada para un segmento específico de la población, por ejemplo, los *datos secundarios* no lograrían cubrir sus objetivos.

En relación a esto, puede considerarse a los usuarios de datos publicados "prisioneros de decisiones sobre la conceptualización realizadas por aquellos que originariamente recogieron los datos" (Jacob, 1994: 345-346).

d) Los *conceptos* también pueden cambiar de significado a lo largo del tiempo. Lo que dificulta la comparación de datos reunidos en distintas fechas.

Dicha comparabilidad también se agrava, cuando unos mismos *conceptos* no se han medido de la misma forma en distintas fuentes de *datos secundarios*.

Pese a estos inconvenientes, el *uso de datos secundarios* se muestra "imprescindible para buena parte de los planteamientos macrosociológicos, en los que se trata de indagar cuestiones referentes a la estructura social global o a la articulación de sus subestructuras" (Beltrán, 1989: 34-35). Por lo que algunos autores se muestran bastante optimistas, y estiman que "el análisis secundario es probable que mantenga una posición dominante en la investigación de las ciencias sociales para el futuro previsible" (Kiecolt y Nathan, 1987: 11).

Indudablemente, a su gran expansión han contribuido los avances habidos en el campo de la informática en los últimos años. Éstos se han materializado, sobre todo, en la creación de *archivos de datos,* y en las posibilidades de tratamiento analítico ahora al alcance de los usuarios de ordenadores personales.

En el Cuadro 6.1 se resumen algunas de las *ventajas* e *inconvenientes* principales de la estrategia de investigación aquí denominada *uso de fuentes documentales y estadísticas*.

6.5. Ejemplo de uso de datos secundarios en una investigación real

Aunque el uso de *datos secundarios* está presente en alguna (o varias) de las fases comprendidas en el desarrollo de cualquier investigación, aquí se ha seleccionado, como ejemplo ilustrativo, una investigación reciente realizada exclusivamente a partir de *fuentes secundarias*. Se trata de la que fuera tesis doctoral del profesor Manuel Justel Calabozo (1995) *La abstención electoral en España, 1977-1993,* publicada por el Centro de Investigaciones Sociológicas (CIS), en la colección *Monografías,* con el número 143.

CUADRO 6.1. Ventajas e inconvenientes del uso de fuentes documentales y estadísticas.

Ventajas	*Inconvenientes*
Precisa menos tiempo y recursos (humanos y económicos) que la investigación con datos primarios.	No disponibilidad, por ser difícil o estar restringido su acceso, o por no adecuarse a los objetivos de la investigación, al período de tiempo o a la población que el analista precisa.
Facilita el acceso a un mayor volumen de información (procedente de amplias y diversas muestras), difícil de alcanzar mediante una sola investigación primaria.	Falta de control y/o conocimiento del proceso de obtención y registro originario de la información. Ello limita la evaluación de la calidad de los datos y su interpretación.
Permite cubrir amplios períodos de tiempo.	Su potencial de favorecer la comparación de datos se restringe cuando acaecen cambios en la organización de la información a lo largo del tiempo.
Ayuda al diseño y realización de un estudio, desde los inicios del proyecto de investigación (la formulación del problema y las hipótesis a comprobar) hasta su conclusión (la interpretación de los resultados).	

En esta investigación se pretende la descripción y explicación de la abstención electoral en España, sus características, factores y evolución, durante el período comprendido entre 1977 y 1993. Concretamente, se limita a dos tipos de procesos electorales: las elecciones generales y las elecciones municipales (a las elecciones autonómicas y europeas apenas se hace referencia).

La inexistencia de información electoral equiparable a la actual impidió alargar hacia atrás el período de estudio de la abstención en España. Por lo que la investigación se circunscribe analíticamente a la etapa democrática actual, fijando su inicio en las elecciones generales del 15 de junio de 1977.

Para cubrir los objetivos del estudio (explicar la abstención electoral y su evolución durante el período analizado), se llevó a cabo:

a) Un *análisis agregado* de los resultados electorales y otros datos estadísticos agregados.
b) Un *análisis individual,* a partir de la información tomada de encuestas realizadas en dichas fechas.

Además, se tomaron datos de otros estudios para comparar los niveles de abstención en elecciones locales y legislativas en la Comunidad Europea. A ello se suma

la correspondiente *revisión bibliográfica* exigida de investigaciones que enmarquen el tema de estudio y los aspectos metodológicos de su análisis.

La explicación ofrecida se refiere tanto a las manifestaciones, como *agregado,* de los electores que no votan (en diferentes ámbitos territoriales o demográficos), como a su dimensión *individual* (en términos de comportamiento político y electoral de los abstencionistas).

Los *datos agregados* de abstención que se analizan en esta investigación proceden, en su mayoría, de los avances de resultados electorales hechos públicos por el Ministerio del Interior. Si bien, la información más precisa (de la facilitada por esta fuente oficial) concierne a los votantes, porque se contabilizan uno a uno en cada escrutinio electoral. En cambio, los abstencionistas resultarán de la diferencia entre el número de electores y el número de votantes (hayan emitido un voto válido, blanco o nulo). Para su cálculo se toman los datos sobre electores procedentes del *Censo electoral,* publicado por el INE.

En su interpretación, téngase presente que esta fuente estadística se ve afectada por la distancia temporal que separa la consulta electoral de la fecha de realización del *censo* o *padrón municipal.* Pese a ello, la información que proporciona la citada *fuente secundaria* fue de gran utilidad para la *descripción* (niveles, fluctuaciones y tendencias) de la abstención en cada proceso electoral, distinguiendo los contextos nacional, regionales, provinciales y locales, y en las zonas rurales y urbanas. Asimismo, posibilitó el establecimiento de relaciones entre niveles de abstención y diferentes características demográficas, económicas y sociales de las provincias y ciudades españolas.

No obstante, el análisis *agregado* efectuado se plantea en términos de *complementariedad* con los datos de *encuesta.* Los *datos de encuesta* reanalizados procedieron, todos ellos, del Centro de Investigaciones Sociológicas (CIS). Es decir, de *encuestas post-electorales* de ámbito nacional, levantadas pocos días después de cada proceso electoral (y con criterios y procedimientos estandarizados).

La *validez* de estas *encuestas* depende, obviamente, de la ausencia de distorsión entre los comportamientos efectivos y los comportamientos confesados en ellas. En realidad, las *encuestas* manejadas por el autor sólo reflejan, aproximadamente, la mitad de los abstencionistas reales (respecto de la fuente oficial consultada). Pese a lo cual, la información se estima de gran interés porque permite:

a) Trazar el perfil personal y contextual de los abstencionistas declarados, y su evolución a lo largo del período de tiempo considerado.
b) Establecer en qué medida los comportamientos individuales vienen determinados (o se explican) por factores medioambientales o individuales.

La información disponible se procesó mediante dos *técnicas analíticas multivariables:* el *análisis factorial* (que agrupó las 14 variables independientes incluidas en cuatro *factores,* explicando el 80% de la *varianza* conjunta), y *el análisis de regresión múltiple* (que mostró ser estadísticamente significativo, aunque la capacidad explicativa de

los modelos obtenidos fue muy reducida: entre el 1 y el 8% de la *varianza* de la variable abstención electoral).

También se realizaron *análisis bivariables* (de cruces de variables) para explorar la relación existente entre cada variable explicativa con la conducta abstencionista. De todo ello se concluyó la propuesta de un modelo integrado por factores subjetivos y objetivos, con el propósito de explicar el comportamiento abstencionista mediante un *path análisis*.

Lecturas complementarias

Dale, A. *et al.* (1988): *Doing secondary analysis*. London, Unwin Hyman.
Glass, G. V. *et al.* (1987): *Meta-analysis in social research*. Beverly Hills, Sage.
González Rodríguez, B. (1994): "La utilización de los datos disponibles". En García Ferrando *et al.* (comp.): *El análisis de la realidad social*. Madrid, Alianza, pp. 245-273.
Kiecolt, K. J. y Nathan, L. A. (1987): *Secondary analysis of survey data*. Beverly Hills, Sage.
Sierra Bravo, R. (1993): *Tesis doctorales y trabajos de investigación científica*. Madrid, Paraninfo.
Stewart, D. M. (1984): *Secondary research: information sources and methods*. Beverly Hills, Sage.
Valles Martínez, M. (1996): *Técnicas cualitativas de investigación social*. Madrid, Síntesis.

Ejercicios Propuestos

1. En un estudio sobre la natalidad, ¿qué fuentes documentales y estadísticas consultaría?
2. Si el estudio fuese sobre el maltrato infantil, ¿a qué fuentes de información secundaria acudiría? Realice una búsqueda bibliográfica e indique las palabras clave que haya utilizado.
3. Describa un estudio en que se haya seguido la estrategia aquí denominada "uso de fuentes documentales y estadísticas". Señale los objetivos del estudio, su diseño y los resultados obtenidos.

7

LA INVESTIGACIÓN SOCIAL MEDIANTE ENCUESTA

De las *estrategias de investigación* existentes, quizás sea la *encuesta* la más conocida y practicada por los investigadores sociales. Sus antecedentes se remontan –como ya se expuso en el Capítulo 1– a los mismos orígenes de la indagación social empírica. En concreto, a estudios ilustres como los realizados en Gran Bretaña por John Sinclair *(Informe estadístico de Escocia,* 1791-1825), James Kay Shuttleworth *(Las condiciones morales y físicas de la vida de los obreros de la industria textil de Manchester,* 1832) y Charles Booth *(Vida y trabajo de los habitantes de Londres,* 1889-1891).

No obstante, es a partir de la Segunda Guerra Mundial cuando la *encuesta* se convierte en la *estrategia* predominante en la investigación social. A ello contribuyeron los impulsos dados, desde la vertiente privada, por los estudios de mercado llevados a cabo, en EEUU por Gallup, Crossley y Ropper; y, desde la vertiente universitaria, por los estudios de Lazarsfeld (sobre todo el que llevó a cabo con Berelson: *La elección del Pueblo,* 1944) y de Stouffer *et al. (El soldado americano,* 1949), principalmente.

7.1. La encuesta como estrategia de investigación

La *encuesta* constituye una *estrategia de investigación* basada en las declaraciones verbales de una población concreta. Puede utilizarse de forma aislada, o en conjunción con otras *estrategias de investigación.* Si bien, sus resultados mejoran cuando en su realización han intervenido otras *estrategias,* ya en fases precedentes (en la confección del *marco teórico* del estudio, el *diseño del cuestionario* y el *diseño muestral)* o en fases posteriores a su desarrollo (en la *validación* e *interpretación* de los resultados de *encuesta).* En el Capítulo 2 se enunciaron algunas de las ventajas e inconvenientes fundamentales de la *triangulación* (o articulación) metodológica. El balance efectuado se inclina a favor de la complementariedad.

7.1.1. Características esenciales de la encuesta

La *encuesta* puede definirse como la aplicación de un *procedimiento estandarizado* para recabar información (oral o escrita) de una *muestra* amplia de sujetos. La *muestra* ha de ser representativa de la *población* de interés; y, la información se limita a la delineada por las preguntas que componen el *cuestionario precodificado,* diseñado al efecto. Entre sus *características* definitorias destacan las siguientes:

a) En la *encuesta* la información se adquiere mediante *observación indirecta*, a través de los respuestas verbales de los sujetos encuestados. Por lo que, siempre cabe la posibilidad de que la información obtenida no refleje la realidad del tema que se investiga. De ahí la necesidad de comprobar la veracidad de los datos recabados.

b) La información abarca un amplio abanico de cuestiones. Puede incluir aspectos *objetivos* (hechos) y *subjetivos* (opiniones, actitudes), del presente o del pasado.

> "Típicamente, las encuestas buscan información sobre la propia conducta y experiencia del individuo, valores y actitudes, características personales y circunstancias sociales. Pero, con frecuencia también buscan información que va más allá del individuo, extendiéndose a sus relaciones, interacciones o actividades con otra gente; proporcionando información sobre contextos sociales, grupos, vecindarios y acontecimientos de los que tienen experiencia" (Hakim, 1994: 52).

c) Para que las respuestas de los sujetos puedan compararse, la *información* se recoge de forma *estructurada.* Se formulan las mismas preguntas, y en el mismo orden, a cada uno de los individuos encuestados.

Pero "ese mismo reduccionismo analítico de la encuesta estadística" es lo que la incapacita "para captar y analizar en profundidad el discurso hablado (relativamente espontáneo y libre) de los sujetos encuestados" (Ortí, 1989: 174).

d) Las respuestas se agrupan y cuantifican para, posteriormente, examinar (mediante *técnicas analíticas estadísticas)* las relaciones entre ellas.

El interés del investigador va más allá del individuo del que obtiene la información. Se extiende a la *población* a la que éste pertenece. De ahí la necesidad de que la *muestra* observada sea representativa de dicha *población.*

e) La *significatividad* de la información proporcionada dependerá de la existencia de *errores de muestreo* (relativos al *diseño muestral* efectuado), y de *errores ajenos al muestreo* (del *diseño del cuestionario,* el *trabajo de campo* y el *tratamiento de los datos* –codificación, grabación, análisis e interpretación–).

En el desarrollo de cualquier *encuesta*, pueden coincidir distintos tipos de *errores*. Éstos cabe resumirlos –en conformidad con lo expuesto por varios autores (Lavrakas, 1987, 1993; Bosch y Torrente, 1993)– en los siguientes:

a) *Errores de muestreo.* Derivan de la *población* que se excluye de la *muestra*, debido a las peculiaridades de la *muestra* diseñada. En concreto, se hallan deter-

minados por: el tamaño de la muestra; cómo se hayan seleccionado las unidades muestrales; y la heterogeneidad de la *población*.

b) *Errores de cobertura*. Referidos a la idoneidad del *marco muestral* elegido para la selección de la *muestra:* si incluye a todos los integrantes de la *población* de interés.

c) *Errores de no respuesta*. Cuando algunas unidades de la muestra no participan, finalmente, en el estudio (por su no localización, negativa personal a participar, u otra causa) o sólo responden a algunas de las preguntas del *cuestionario*. Por ejemplo, las personas de clase alta, o que residen en viviendas de difícil o imposible acceso a los encuestadores, suelen estar infrarrepresentadas en las *encuestas*.

La importancia de este tipo de *error* dependerá de la magnitud de la "no respuesta". También influye el perfil de las personas que responden al *cuestionario:* si éste difiere o no de aquellos que deciden eludir respuestas.

d) *Errores de medición*. No todos los datos que se obtienen por *cuestionario* suponen medidas adecuadas de los conceptos que se analizan. Estas inadecuaciones pueden deberse a:

1) *Errores* en el *diseño del cuestionario* (formulación y orden de las preguntas), y en el modo de administrarlo (mediante entrevista personal, telefónica o por correo).
2) *Sesgos* introducidos por el *entrevistador* en el desarrollo de la entrevista (si la encuesta no es autoadministrada).
3) *Sesgos* atribuibles al *entrevistado,* cuando no aporta respuestas veraces a las preguntas que se le formulan (debido a fallos de memoria, o debido a factores de la situación social de entrevista, como la conocida tendencia de ajustarse a las respuestas deseables socialmente).

7.1.2. Ventajas e inconvenientes de la encuesta

La popularidad que ha adquirido la *encuesta* en la investigación social se debe, sin duda, a las muchas *posibilidades* que ofrece al investigador. Comparada con otras estrategias de investigación, la *encuesta:*

a) Resalta por su adecuación para obtener información diversa, de un conjunto amplio de personas, ubicadas en distintas áreas geográficas.

Aunque elevado, el coste de la investigación (en tiempo y dinero) resulta inferior al necesario para la materialización de otras estrategias de investigación (en una población similar).

b) Se distingue por su elevado grado de *fiabilidad*. Si bien, hay que matizar que la *fiabilidad* estará condicionada al tipo y magnitud de los *errores* cometidos en su realización (a los que se ha hecho mención en el subapartado 7.1.1).

De acuerdo con Mayntz *et al.* (1983), la *fiabilidad* se afianza gracias, sobre todo, a la *estandarización* de las respuestas (a mayor estandarización, mayor fiabilidad), y a la formulación de las preguntas (la fiabilidad crece conforme au-

menta la claridad expositiva de las preguntas). En cambio, la *validez* se ve afectada por la situación de entrevista característica de la *encuesta* (inadecuada para abordar determinadas cuestiones). También puede verse afectada por la formulación de las preguntas (dependiendo de si éstas constituyen, realmente, *indicadores* válidos de los *conceptos* que se pretenden medir). En el Capítulo 4 (apartado 4.5) ya se han tratado las cuestiones básicas de *validez* y *fiabilidad* inherentes a todo proceso de *medición*.

c) La *encuesta* permite la comparación de datos obtenidos en estudios realizados en fechas, países o áreas geográficas diferentes. Pero, para ello, es preciso que se utilicen las mismas preguntas en todas las encuestas.

Para una comparabilidad máxima, las preguntas deben estar redactadas y ordenadas en el *cuestionario* de manera equiparable; y las características técnicas fundamentales de las encuestas deben permanecer constantes.

La utilidad de la *encuesta* aumenta "cuanto mayor sea la posibilidad de poder comparar los datos (obtenidos) con otros similares y anteriores en el tiempo o procedentes simultáneamente de otras poblaciones" (López Pintor, 1989: 360).

d) La aplicación de la *teoría de la probabilidad* y del *muestreo* permite el cálculo de la *significatividad estadística,* dando basamento matemático a la generalización de los datos de *encuesta*.

Pese a sus ventajas, la *encuesta* también se enfrenta a una serie de *obstáculos,* que impiden su autosuficiencia. Éstos pueden resumirse en los siguientes:

a) La aplicación de la *encuesta* resulta desaconsejable en poblaciones con dificultades para la comunicación verbal (niños pequeños, personas analfabetas o con escasos recursos lingüísticos).

b) La información que se obtiene se halla condicionada por la formulación de las preguntas, y por la veracidad de las respuestas de los encuestados.

Las *preguntas* son, en su mayoría, *cerradas*. Por esta razón, las opciones de respuesta se limitan a las previamente especificadas por el investigador. De ahí la conveniencia de complementar los datos de *encuesta* con la información reunida mediante otras *estrategias de investigación*.

El *estudio de casos,* por ejemplo, puede ayudar (mediante la *observación* y las *entrevistas informales,* individuales o grupales) a la *validación* e *interpretación* de los datos estadísticos de *encuesta*.

La comprensión de las relaciones estadísticas (de las variables observadas en una *encuesta)* precisa, con frecuencia, de información sobre las experiencias vitales de las personas encuestadas. Esta información es difícil de generar mediante la *encuesta*. Por el contrario, las *técnicas cualitativas* de obtención de datos resultan más apropiadas. Aunque, también éstas presentan deficiencias, generalmente relativas al habitual pequeño número de casos observados en los *estudios cualitativos*. Ello genera un continuo debate sobre la *representatividad* de las conclusiones de las *investigaciones* exclusivamente *cualitativas*.

c) La presencia del entrevistador (cuando la *encuesta* se hace mediante entrevista personal o telefónica) provoca *efectos* de carácter *reactivo* en las respuestas de los entrevistados.

Los problemas de *reactividad* pueden, sin embargo, reducirse con un buen diseño del cuestionario, junto a una adecuada selección y formación de los entrevistadores –como se verá en los apartados siguientes–.

d) Al ser la mayoría de las *encuestas transversales* (son minoría las *encuestas panel* o encuestación repetida a un mismo grupo de personas) y carentes del *control experimental* (control *a priori),* pueden confundirse simples correlaciones entre variables con verdaderas relaciones causales.

No obstante, el uso de *técnicas estadísticas multivariables* ayuda a la reducción de esta inexactitud.

e) La realización de una *encuesta* precisa de la organización de un *trabajo de campo* complejo y costoso. Especialmente, si se desea abarcar, mediante *entrevista personal,* segmentos amplios y dispersos de la población.

Téngase en cuenta que el precio estándar por entrevista puede llegar a oscilar (a mediados de los años noventa) entre 4.000 y 6.000 pesetas, dependiendo de la complejidad del *cuestionario* y del procedimiento empleado para la *selección muestral.* Ello limita su uso a quien pueda económicamente financiarlas.

En el Cuadro 7.1 se resumen algunas de las ventajas e inconvenientes de la *encuesta* como *estrategia de investigación.*

CUADRO 7.1. Ventajas e inconvenientes de la encuesta.

Ventajas	*Inconvenientes*
Permite abarcar un amplio abanico de cuestiones en un mismo estudio.	No resulta adecuada para el estudio de poblaciones con dificultades para la comunicación verbal.
Facilita la comparación de resultados (al basarse en la estandarización y cuantificación de las respuestas).	La información se restringe a la proporcionada por el individuo (a preguntas generalmente cerradas).
Los resultados del estudio pueden generalizarse, dentro de los límites marcados por el diseño muestral efectuado.	La presencia del entrevistador provoca efectos reactivos en las respuestas.
Posibilita la obtención de una información significativa, cuando no acontezcan graves errores en su realización.	La carencia de referencias contextuales y vitales de los individuos limita la interpretación de los datos de encuesta.
Ventaja económica: puede obtenerse un volumen importante de información a un mínimo coste (económico y temporal).	Acusa imprecisión para el estudio de la causalidad.
	La existencia de obstáculos físicos (edificios vigilados, porteros automáticos, contestadores automáticos) dificultan el contacto con las unidades muestrales.
	El desarrollo de una encuesta amplia resulta complejo y costoso (sobre todo en encuestas personales).

7.1.3. Modalidades de encuesta

Aparte de la distinción, anteriormente referida, entre *encuesta transversal* y *encuesta panel,* existen tres modalidades principales de *encuesta,* en función de cómo se administre el *cuestionario:* mediante entrevista personal, telefónica, o por correo (como variedad de *encuesta autoadministrada).* Cada una de estos tipos de *encuesta* lleva consigo unas ventajas, pero también unos inconvenientes, que deberán sopesarse a la hora de escoger entre las distintas modalidades de *encuesta.*

En la decisión sobre qué variedad de encuesta elegir se valora, sobre todo, su adecuación al logro de los objetivos de la investigación. En especial:

a) El *tema* o temas que se abordan.
b) La amplitud y la complejidad del *cuestionario* que se precise.
c) La *población* de interés que forme el *universo* del estudio.
d) El *tiempo* concedido para su realización.
e) Los *recursos* (económicos y humanos) disponibles.

• *Encuesta personal o entrevista "cara a cara"*

Hasta ahora, ésta constituye la modalidad de *encuesta* más practicada en la investigación social. Consiste en la administración del *cuestionario* mediante *entrevista personal,* por separado, a cada uno de los individuos seleccionados en la *muestra.* Es el entrevistador quien formula las preguntas y quien anota las respuestas en el *cuestionario.* Por esta razón, el éxito de la investigación descansa bastante en la actuación del entrevistador.

El *entrevistador* no sólo debe de tener la capacidad de establecer empatía con el entrevistado y conseguir su cooperación, sino que también debe de estar adecuadamente preparado para:

a) Formular correctamente las preguntas del *cuestionario.*
b) Asegurar la adecuación de las respuestas y su correspondiente anotación.
c) Tomar decisiones en el campo, sin la asistencia de un *supervisor.*

En este tipo de *encuesta,* la intervención del entrevistador se convierte en decisiva. Puede facilitar el desarrollo del *trabajo de campo,* pero también puede obstaculizarlo. Su presencia permite:

a) El tratamiento de temas complejos. El entrevistador puede aclarar cuestiones no entendidas. También puede ofrecer *ayudas visuales* (tarjetas de respuestas, gráficos) para la comprensión de preguntas complejas.
b) *Comprobar* la comprensión de las preguntas y la consistencia de las respuestas.

c) Recoger información suplementaria, sobre características personales del entrevistado, de su entorno sociofamiliar, e incluso de la vivienda donde habita (cuando la entrevista se lleva a cabo en el domicilio del entrevistado). Esta información complementará los datos recogidos en el *cuestionario*.
d) Despertar el interés del entrevistado, motivándole para responder (con sinceridad y claridad) las cuestiones que se le pregunten. Ello contribuye a que ésta sea la modalidad de encuesta que logra un mayor porcentaje de respuestas.

Pero, la presencia del entrevistador también puede provocar problemas importantes en el desarrollo de la *encuesta*. Éstos pueden agruparse en tres *problemas* fundamentales:

a) *De acceso* a las viviendas particulares, debido a la llamada "inseguridad ciudadana" (más sentida en las grandes áreas urbanas).
b) *De localización* de determinados grupos de población. Sobre todo, si el *trabajo de campo* se realiza por la mañana, durante el horario laboral.
A menos que el entrevistador concierte cita previa (o acuda al domicilio a horas diferentes), habrá una sobrerepresentación de amas de casa, jubilados y parados en la *muestra* finalmente observada.
El acceso a grupos de población menos "accesibles" lógicamente complica y encarece el *trabajo de campo*.
c) *De reactividad* del encuestado, que puede afectar al contenido de las respuestas. Así, por ejemplo, la elección de respuestas consideradas "socialmente deseables" es más usual en *entrevistas personales* que en las otras modalidades de *encuesta*.

Dada la relevancia que adquiere la actuación del *entrevistador* para los resultados de la investigación, su trabajo ha de supervisarse. La existencia de un equipo de *supervisores* resulta imprescindible en cualquier encuesta cumplimentada mediante entrevista (personal o telefónica). Éstos han de controlar la mediación del entrevistador, tanto en la selección de las unidades muestrales (que las personas entrevistadas se correspondan a la *muestra* diseñada), como durante el proceso de entrevista. También deben controlar que el *trabajo de campo* se haga con prontitud y efectividad.

La necesidad de numeroso personal (entrevistadores y supervisores), encarece los costes de la investigación. A ello se suma el tiempo preciso para su ejecución. Todo lo cual obstaculiza la práctica de esta modalidad de *encuesta*.

En el Cuadro 7.2, se resumen las ventajas e inconvenientes principales de la *encuesta personal*.

CUADRO 7.2. Ventajas e inconvenientes de la encuesta personal.

Ventajas	*Inconvenientes*
Permite la consecución de un mayor porcentaje de respuestas.	Encarece los costes del estudio, en tiempo y dinero.
Favorece el tratamiento de temas complejos.	Dificultad para acceder a domicilios particulares y a determinados grupos de población.
Se obtienen respuestas de mayor calidad y espontaneidad.	La presencia del entrevistador puede provocar efectos reactivos en las respuestas de los entrevistados.
El entrevistador puede recabar información complementaria del entrevistado ajena al cuestionario.	

• *Encuesta telefónica*

A diferencia de la modalidad anterior, en la *encuesta telefónica* la comunicación entrevistador-entrevistado acontece a través del hilo telefónico. Ello supone una serie de *ventajas* importantes:

a) Abarata los costes del *trabajo de campo*. Aunque el coste de la llamada telefónica es elevado, el presupuesto necesario para una *encuesta telefónica* es sensiblemente inferior a una *encuesta personal*. Precisa menos personal (entrevistadores, supervisores), y se elimina la partida presupuestaria destinada a sufragar los costes de desplazamiento de los entrevistadores.

b) Acorta el tiempo de realización del *trabajo de campo*. Desde una misma central, un reducido número de entrevistadores pueden, en un mismo día, efectuar una cifra elevada de entrevistas (imposible de alcanzar mediante *entrevista personal)*. La duración de cada entrevista también suele ser breve (entre 10 y 15 minutos).

c) Permite abarcar núcleos dispersos de población, sin apenas coste (en tiempo y dinero). Al no tener que desplazar entrevistadores, no hay razón para agrupar a los individuos en unos *puntos de muestreo* concretos. Puede alcanzarse, en cambio, el ideal de dispersión muestral.

El teléfono facilita el contacto con sujetos de cualquier localidad geográfica, casi al instante, y a un mínimo coste.

d) Puede accederse a los grupos de población menos "accesibles" (bien por su profesión, o bien por las características de las viviendas o entornos donde residen –edificios vigilados, porteros automáticos, viviendas retiradas o en barrios marginales–).

e) Facilita la repetición de los intentos de selección de las *unidades muestrales*. Sin duda, resulta más cómodo y económico llamar reiteradamente a un mismo nú-

mero de teléfono (hasta localizar a alguien en el domicilio), que enviar a un entrevistador. Esto permite aumentar el número de intentos de selección, antes de reemplazar una unidad de la muestra.

f) Posibilita la *supervisión* de las entrevistas durante su realización, al reunir en un mismo edificio a entrevistadores y supervisores (o personal a cargo de la investigación). Lo que permite la consulta de cualquier contrariedad que aparezca durante el *trabajo de campo*.

Asimismo, el disponer de programas de ordenador específicos para la realización de *entrevistas telefónicas* (sistemas CATI: *Computer Assisted Telephone Interview),* permite un mayor control de la actuación de los entrevistadores.

En la actualidad, el uso de *entrevistas telefónicas* asistidas por ordenador (CATI) –que comenzaron a implantarse en EEUU durante los años setenta– se halla muy extendido. Consiste en un programa informático de entrevista, en el que se introducen las preguntas del *cuestionario* (en el orden especificado), junto con las respuestas (si la pregunta es *cerrada).* En todo caso, para cualquier pregunta debe especificarse: el tipo de pregunta, el tipo de respuesta (numérica, alfanumérica), y el tamaño de la respuesta. Además, debe también determinarse dónde comienza y termina el texto, y dónde se almacenan los datos.

Las preguntas aparecen reflejadas en la pantalla del ordenador. El entrevistador las lee, e introduce en el ordenador las respuestas de los entrevistados. De esta forma se reducen bastante los costes de la investigación (en tiempo y dinero). Se elimina el trabajo tedioso de *grabación* posterior (en el ordenador) de los datos anotados en cientos o miles de *cuestionarios de papel.* Este trabajo es generalmente efectuado por personal especializado, aunque no está exento de errores en la grabación de las respuestas.

La aplicación del programa CATI también aumenta la calidad de los datos reunidos. Precisamente, porque el programa permite –como indica Saris (1994: 170)– "limpiar los datos mientras el encuestado está todavía disponible". Tan pronto como se emite la respuesta a una pregunta, el entrevistador puede contrastar dicha información con las respuestas anteriores. De esta forma comprueba la consistencia y calidad de las respuestas.

Sin embargo, existen –siguiendo a este mismo autor– tres *inconvenientes* esenciales en los distintos procedimientos de recogida de información mediante ordenador (CADAC: *Computer Assisted of Data Collection):*

a) El tamaño de la pantalla del ordenador es inferior a la página de un *cuestionario* impreso en papel. Lo que dificulta su lectura y la contextualización de la pregunta.

b) Los *cuestionarios de papel* proporcionan una visión continua de las preguntas y de las respuestas. Por lo que puede verse, con mayor facilidad, dónde se han cometido errores.

c) En los procedimientos CADAC, se tiene que pulsar el teclado del ordenador y, al mismo tiempo, mirar la pantalla para comprobar los datos que se han entrado. Si no se mira la pantalla, mientras se teclea, pueden introducirse errores importantes en el registro de las respuestas. De ahí la conveniencia de utilizar *pantallas de resumen.*

El tener que estar continuamente mirando al teclado y a la pantalla, puede también alargar la duración de la entrevista (respecto al método tradicional de papel y lápiz. A ello se suma, el usual pequeño espacio dado al entrevistador para que registre anotaciones relacionadas con el desarrollo de la entrevista (Saris, 1994).

Los sistemas CADAC incluyen tanto la *encuesta telefónica* (CATI), como la *entrevista personal* (CAPI) y la *encuesta autoadministrada* (CSAQ):

- En la *entrevista personal asistida por ordenador* (CAPI), el entrevistador se acompaña de un ordenador portátil para la realización de las entrevistas. Su uso es similar al sistema CATI.
- En el *cuestionario autoadministrado asistido por ordenador* (CSAQ), el entrevistador deja la lectura de las preguntas y la entrada de las respuestas al encuestado.

Estos dos programas informáticos (CAPI y CSAQ) son de más reciente implantación, y de uso más restringido, que el CATI. Ello se debe, sobre todo, al elevado número de ordenadores portátiles que precisa su puesta en práctica. Ordenadores cuyo precio (pese a la continua reducción), todavía limita su uso mayor.

Además de facilitar la recogida de información, el ordenador ayuda a la selección de las *unidades muestrales*. En una *encuesta telefónica*, los números de teléfono pueden obtenerse, laboriosamente, mediante su extracción aleatoria de las guías telefónicas. Esta tarea se resuelve, cómodamente, con la ayuda del programa de ordenador que genera, al azar, números de teléfono.

Este método de selección muestral resulta más cómodo y rápido que el tradicional. Sin embargo, continúa sin resolverse un problema importante: la llamada innecesaria a una residencia inexistente. Se desconoce el lugar a donde se llama. Hecho que determina que no pueda precisarse siempre la razón de la "no respuesta". Ésta puede incluso alcanzar un porcentaje superior al obtenido mediante el método de selección muestral tradicional.

De las críticas habituales a la *encuesta telefónica,* la más frecuente concierne a *errores de cobertura:* se excluye a aquellos que carecen de teléfono o cuyos números no figuran en el *marco muestral* utilizado. Aunque cada vez son más las viviendas que disponen de teléfono (el 76% de los hogares, según estimaciones de Bosch y Torrente, 1993), continúan observándose diferencias por hábitat y estatus social. La cobertura telefónica es inferior en las capas bajas de la sociedad y en entornos rurales. Por esta razón, estos grupos de población no obtienen la debida representación en la *muestra.* Ello dificulta la generalización de los resultados de la investigación al conjunto de la *población.*

A estas *críticas* se suman otras, relacionadas con las características del medio a través del que se desarrolla la comunicación entrevistador-entrevistado:

a) Como la comunicación es únicamente verbal, el entrevistador no puede recurrir a *ayudas visuales*. Por lo que deberá demostrar mayores habilidades persuasivas y de conversación.

b) Cuando el entrevistador dicta distintas respuestas a las preguntas (y estas respuestas son variadas), se exige al entrevistado el ejercicio constante de la memoria. Esto provoca un problema importante: que el entrevistado no medite la pregunta. Como Bosch y Torrente (1993: 19) observan, "debido a la limitada capacidad memorística y a la presión del tiempo, existe la tendencia a responder la primera cosa que viene a la mente".

También existen obstáculos físicos, como líneas constantemente ocupadas, contestadores automáticos, números de teléfonos desconectados. Estos "obstáculos", además de imposibilitar la comunicación entrevistador-posible entrevistado, inflan el porcentaje de "no respuesta" atribuido a la *encuesta telefónica*. Pese a ello, el teléfono se presenta, en la actualidad, como un medio fácil y rápido para contactar al entrevistado (y a cualquier hora del día). A ello se suma la ventaja de mostrarse como un medio menos "inhibidor" que la entrevista "cara a cara". Lo que ayuda a la sinceridad de las respuestas.

En el Cuadro 7.3 figuran ventajas e inconvenientes esenciales de la *encuesta telefónica*.

CUADRO 7.3. Ventajas e inconvenientes de la encuesta telefónica.

Ventajas	*Inconvenientes*
Reduce el coste y el tiempo de realización del trabajo de campo.	Errores de cobertura, al excluirse a personas carentes de teléfono.
Facilita el acceso a domicilios particulares, y la repetición de los intentos de selección.	Existencia de obstáculos físicos, que dificultan el contacto con las unidades muestrales: contestadores automáticos, líneas ocupadas, teléfonos desconectados.
Posibilita la inclusión en la muestra de núcleos de población dispersos y de personas de difícil localización.	Imposibilidad de recurrir a ayudas visuales para la cumplimentación del cuestionario.
Permite la supervisión durante la realización de las entrevistas.	Exige una mayor capacidad de comunicación entre el entrevistador y el entrevistado.
Inhibe menos que la entrevista personal.	Demanda del entrevistado una mayor capacidad memorística (de retención de preguntas y respuestas).
	El entrevistador no puede recabar información suplementaria del entrevistado.
	La duración de la entrevista suele ser menor. Lo que supone la reducción del cuestionario.

A estas ventajas e inconvenientes hay que añadir las comunes a cualquier administración del cuestionario mediante entrevistador. A ellas se hizo referencia en la exposición de la *encuesta personal.*

• *Encuesta por correo*

Se engloba dentro de la categoría genérica de "encuesta autoadministrada". Ésta comprende cualquier tipo de sondeo de opinión que se caracterice por ser el propio encuestado quien lee el *cuestionario* y anota las respuestas.

En su realización el encuestado puede estar o no acompañado de algún responsable de la investigación (un entrevistador u otra persona). Un ejemplo de la primera situación es cuando se encuesta a estudiantes "recluidos" en una aula de un centro de enseñanza. Éstos pueden consultar cualquier duda sobre el *cuestionario* al personal presente en el aula. En cambio, la *encuesta por correo* se distingue por la total ausencia de entrevistadores u otros integrantes del equipo investigador. Es el mismo encuestado quien rellena el cuestionario y quien lo remite (por correo) a la empresa o centro que se lo ha enviado. Esto permite:

a) Ampliar la cobertura de la investigación. Alcanzar áreas aisladas, y a aquellos miembros de la población a quienes los entrevistadores encontrarían difícil de localizar en sus domicilios.

b) Abaratar los costes del *trabajo de campo.* Se eliminan los gastos destinados a entrevistadores y supervisores; al igual que las partidas para sufragar gastos de desplazamiento.

Los gastos del *trabajo de campo* se concentran en los envíos postales (sobres, sellos), que suponen menos costes que los de otras modalidades de *encuesta.* Apenas exceden de un tercio de lo que costaría una *encuesta personal* de características similares. Esto permite aumentar el tamaño inicial de la *muestra,* sin apenas suponer incrementos importantes en el coste de la investigación.

c) Evita el *sesgo en las respuestas* que pudiera producir la presencia del entrevistador. El sentimiento de privacidad y de anonimato, que proporciona la encuesta por correo, la convierte en el medio más adecuado para tratar temas "delicados" (de la conducta o actitudes de las personas).

d) Ofrece al encuestado más tiempo para reflexionar sus respuestas; y la posibilidad de poder consultar a otras personas, o cualquier documento que estime necesario. Por esta razón, la *encuesta por correo* se muestra como la más pertinente cuando se precisa información detallada (que exija una mayor reflexión de las respuestas); pero, no cuando se deseen respuestas espontáneas.

Pese a sus grandes ventajas, la *encuesta por correo* es la menos practicada en la investigación social. A ello contribuyen los siguientes *inconvenientes:*

a) La proporción de encuestados que remiten el *cuestionario* (y debidamente cumplimentado) es bastante inferior al porcentaje de respuesta obtenido por *entrevista personal* o *telefónica*. Lo que es más, la *muestra* final puede estar sesgada. Las personas que contestaron el *cuestionario* pueden presentar un perfil sociodemográfico bastante diferente del de aquellas que optaron por no cumplimentarlo.

Sin embargo, el *porcentaje de respuesta* puede aumentarse:

1) Con un buen *diseño del cuestionario:* no muy extenso, atractivo, y fácil de rellenar.
2) Incluyendo una *carta de presentación*. Ésta ha de destacar la importancia de la cooperación del destinatario en el estudio. También ha de garantizar el anonimato de la información que se proporcione.
3) Adjuntando un sobre con *contrareembolso* para facilitar la remisión del *cuestionario*.
4) Enviando continuos *recordatorios* (que incluyan copia del *cuestionario)* a aquellos que, pasado un determinado período de tiempo (al menos 15 días desde la entrega del *cuestionario),* no lo hayan aún remitido.

b) La imposibilidad de *controlar* si fue la persona inicialmente seleccionada quien, en realidad, rellena el *cuestionario;* si lo hace sola o con la ayuda de otras personas; si el *cuestionario* se cumplimenta en un ambiente tranquilo o la atención del encuestado es, por el contrario, continuamente interrumpida. Todos estos factores cuestionan la *validez* de las *respuestas*.

c) El encuestado puede leer todo el *cuestionario* antes de rellenarlo. Esto limita la eficacia de las *preguntas de control,* y de cualquier acercamiento progresivo a determinadas cuestiones *(técnica del embudo)*.

d) La dificultad de *asistir* al individuo para que estructure sus respuestas o comprenda términos complejos. Esto limita su uso con personas de escaso nivel educativo.

A estos inconvenientes se suma la exigencia de que el *marco muestral* esté lo más actualizado y completo posible. Aunque los *cuestionarios* puedan enviarse sin especificar el nombre del destinatario, se recomienda su personalización.

En el Cuadro 7.4 se resumen ventajas e inconvenientes observados en la *encuesta por correo*.

CUADRO 7.4. Ventajas e inconvenientes de la encuesta por correo.

Ventajas	*Inconvenientes*
Alcanza áreas aisladas y a personas de difícil localización.	Elevado porcentaje de no-respuesta.
Abarata los costes del trabajo de campo.	Inasistencia al encuestado para clarificar y motivarle a responder las preguntas del cuestionario.
Reduce el sesgo en las respuestas debido a la presencia del entrevistador.	La persona seleccionada puede no ser quien rellena el cuestionario.
Ofrece privacidad para responder el cuestionario.	Puede leerse todo el cuestionario antes de cumplimentarlo, lo que limita la eficacia de las preguntas de control y de la técnica del embudo.
El encuestado dispone de más tiempo para reflexionar sus respuestas y comprobar información.	

7.1.4. Fases de la encuesta

En la realización de una *encuesta* convergen diferentes fases. Desde la formulación y delimitación de los *objetivos* específicos de la investigación; hasta el *diseño de la muestra,* la elaboración del *cuestionario,* la preparación y realización del *trabajo de campo;* para finalizar con el *tratamiento y análisis* de la información recabada. En la Figura 7.1 se esquematizan las fases básicas en la realización de una *encuesta.*

De las fases señaladas en el gráfico, las iniciales deciden el buen funcionamiento de la investigación. Sin duda, la estructuración y éxito final de la *encuesta* depende, en gran medida, de la adecuación del *cuestionario:* primero, a los objetivos específicos de la investigación; y, segundo, a las características de la población que se analiza. Cuanto más claros estén los *objetivos* del estudio, más fácil será la traducción de *conceptos* a preguntas concretas y pertinentes.

Asimismo, la acotación de la *población* (niños, estudiantes, amas de casa, jubilados, profesionales) orientará tanto el *diseño* de la *muestra* como del *cuestionario.*

Nada de esto puede hacerse con propiedad sin la previa y necesaria *consulta bibliográfica.* La lectura de investigaciones teórico-prácticas constituye un buen punto de partida, que orientará y llevará al investigador, primero, a precisar qué quiere analizar *(objetivos del estudio),* y cómo lo va a efectuar *(modalidad de encuesta).* Después, en función de estos dos aspectos claves, diseñará la *muestra* y el *cuestionario.* Todo ello estará.

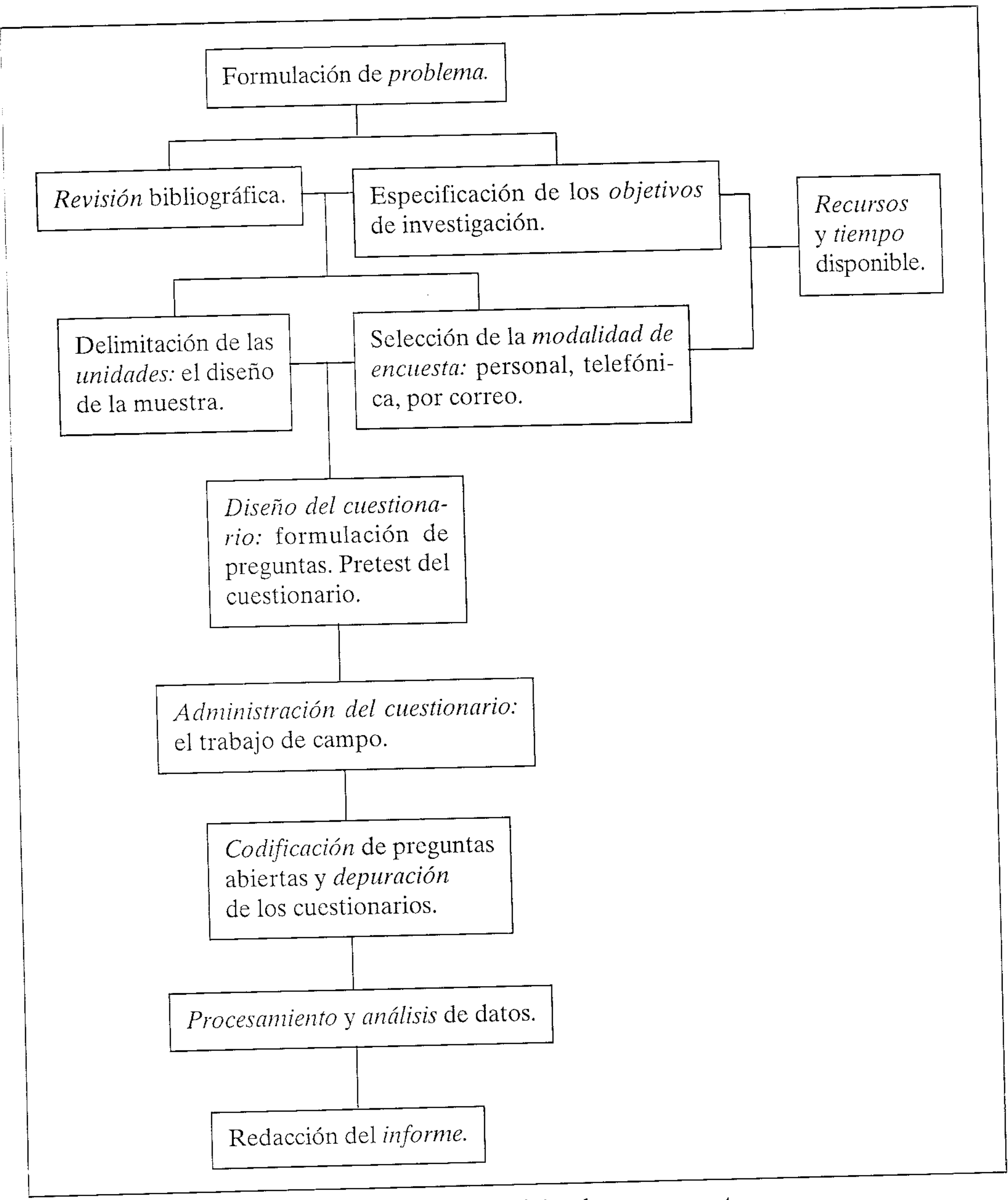


Figura 7.1. Fases esenciales de una encuesta.

a su vez, determinado por los *recursos* (económicos, materiales y humanos), y el *tiempo* que el investigador disponga para la realización de la investigación.

Las limitaciones presupuestarias y temporales marcan la elección de la *modalidad de encuesta* (personal, telefónica, por correo u otra variedad de encuesta autoadmi-

nistrada). En función de cuál haya sido la modalidad finalmente seleccionada, se diseña la *muestra* y se confecciona el *cuestionario*.

En los capítulos precedentes, se han tratado las fases previas comunes a cualquier investigación. En los siguientes apartados se expondrán las fases específicas a la *encuesta* (el *diseño del cuestionario* y el *trabajo de campo)*. Para las fases posteriores (el *análisis de los datos* y la *redacción del informe)*, se remite a los Capítulos 9 y 11, respectivamente.

7.2. El diseño del cuestionario

En la investigación mediante *encuesta*, el instrumento básico para la recogida de información lo constituye el *cuestionario* (estandarizado). Pero el *cuestionario* no es una técnica de obtención de datos exclusiva de la *encuesta*. Puede también aplicarse en otras *estrategias de investigación* (como el *experimento* o el *uso de fuentes secundarias* –para el vaciado de la información contenida en informes o expedientes, por ejemplo–), de forma aislada, o complementando otras *técnicas* de recogida de información.

El *cuestionario* consiste en un listado de *preguntas estandarizadas* (leídas literalmente y siguiendo el mismo orden al entrevistar a cada encuestado). Su formulación es idéntica para cada encuestado.

7.2.1. Tipos de preguntas

Existe una gran variedad de preguntas, pues son varios los criterios de clasificación posibles. El más usual diferencia las *preguntas cerradas* (o precodificadas) de las *abiertas*.

- *Preguntas cerradas*

Las *preguntas cerradas* (también denominadas *precodificadas* o de *respuesta fija)* son aquéllas cuyas respuestas ya están acotadas, cuando se diseña el *cuestionario*.

El investigador determina, previamente, cuáles son las diversas opciones de respuesta posibles. El encuestado se limita a señalar cuál, o cuáles (si la pregunta es *múltiple)*, de las opciones dadas refleja su opinión o situación personal.

Las respuestas se listan verticalmente. A cada una de ellas se le adjunta un número. Ese número constituye el *código numérico,* que facilitará tanto la transferencia de la respuesta verbal a un *fichero de datos* informatizado, como su posterior análisis estadístico. Ello repercute en una mayor exigencia de rigor y exhaustividad en su *formulación:*

a) El investigador deberá documentarse (con anterioridad a la redacción de la pregunta) sobre las distintas alternativas de respuesta existentes a la cuestión que se pregunta.

Las diferentes *categorías* u opciones de respuesta deberán, asimismo, cumplir los requisitos de *exhaustividad, precisión* y *exclusión* señalados en el Capítulo 4, a cuyo repaso remito.

Si el investigador prevé la posibilidad de que existan otras opciones de respuesta (diferentes a las dadas), deberá incluir la opción "otros". A esta opción se le dotará de espacio suficiente que facilite su especificación.

EJEMPLOS DE PREGUNTA CERRADA CON LA OPCIÓN "OTROS"

Podría decirme: ¿con quién vive Ud.?

Solo	1
Con su cónyuge	2
Con su cónyuge e hijos	3
Con sus hijos	4
Con otros familiares	5
Otra situación (especificar).......	
....	

¿Cómo recibe su retribución familiar?

Asignación mensual fija	1
Por trabajo realizado	2
Me dan lo que necesito	3
Me dan lo que pueden	4
Otra forma ¿cuál?......................	
....	
....	

Respecto a las opciones "no sabe", "no contesta", se aconseja su no inclusión expresa en la pregunta, salvo que se estime de interés en su formulación. La experiencia muestra que éstas constituyen opciones de respuesta muy recurridas, cuando el encuestado no quiere pensar o manifestar una respuesta concreta. Este problema se evidencia más en *cuestionarios autoadministrados,* cuyo cumplimiento se deja al arbitrio del encuestado.

En caso que se opte por esta recomendación, se instruye al personal encargado de la *grabación* de los *cuestionarios* (ya completados) a que introduzca, automáticamente, el *código* correspondiente al "no contesta", a aquellas preguntas sin respuesta. Este *código* suele ser el número 9 ó el 0, si sólo se precisa de un dígito para la *grabación* de las respuestas. Si se precisasen dos dígitos, puede optarse entre el 99 o el 00.

Bourque y Clark (1994) matizan que la exclusión de las categorías de respuesta "no sabe", "no contesta", adquiere mayor relieve en indagaciones conductuales o de hecho. Por el contrario, cuando la fuente de información son *documentos* existentes u *observaciones* registradas, aconsejan la inclusión de dichas categorías de respuesta. En su opinión, "uno de los mayores errores que se cometen al extraer datos de informes es no advertir que se buscó información y no se encontró" (Bourque y Clark, 1994: 19).

Por su parte, Converse y Presser (1994) insisten en la recomendación de no proporcionar una categoría de respuesta intermedia. Estos autores advierten que, cuando se ofrece una alternativa de respuesta intermedia, el 20% de los encuestados la escogen, pese a no ser ésta la alternativa que habrían elegido, si no se hubiera ofrecido en el enunciado de la pregunta.

b) El investigador también deberá especificar si la pregunta formulada admite una única respuesta o varias *(pregunta múltiple)*. En este último caso, deberán darse instrucciones expresas sobre cuántas opciones de respuesta se admiten (si se establecen límites de cantidad) y la manera de indicarlas.

EJEMPLOS DE PREGUNTA CERRADA MÚLTIPLE

¿A qué espectáculos artísticos le gusta asistir? Indique los tres que prefiere (pregunta múltiple)

Cine	1
Teatro	2
Opera	3
Concierto	4
Zarzuela	5
Revista musical	6
Circo	7
Otros (especificar)...................	
....	

¿Qué tipo de pensión percibe Ud.? (posible pregunta múltiple)

De la Seguridad Social	1
Clases pasivas	2
Asistenciales, de un organismo público	3
Benéficas (CARITAS)	4
Privadas (sociedades médicas, financieras, empresas)	5
Otros (especificar)	
....	

Cuando la pregunta incluye muchas alternativas de respuesta, en las entrevistas "cara a cara" es habitual el recurso a *tarjetas*. En vez de leer las distintas opciones de respuesta (y forzar al entrevistado a su memorización), se le entrega una *tarjeta* para que visualice las distintas opciones posibles.

EJEMPLOS DE PREGUNTA MÚLTIPLE CON TARJETA

De las siguientes maneras de invertir dinero, ¿cuáles prefiere Ud.? (respuesta múltiple) (mostrar *tarjeta A)*

¿Qué cualidades, de las siguientes, admira Ud. más en una persona? (respuesta múltiple). Señalar sólo tres *(Tarjeta B)*

En cuenta corriente	1	Lealtad	01
En libreta a plazo fijo	2	Respeto	02
En Deuda Pública o en Bonos del Estado	3	Sinceridad	03
Invertir en Bolsa	4	Tolerancia	04
En fondo de pensiones	5	Amabilidad	05
Compra de vivienda	6	Simpatía	06
Compra de joyas u obras de arte	7	Entrega	07
Otras (especificar)...............		Perseverancia	08
....		Humildad	09
....		Comprensión	10
		Sociabilidad	11
		Sencillez	12

c) En la *codificación* de las *respuestas*, deberían seguirse las mismas pautas en todas las preguntas del *cuestionario*. Por *ejemplo,* codificar siempre las respuestas "Sí", como "1"; "No", como "2"; y "No contesta", como "0".

También se estima conveniente que el *código* asignado se corresponda con el significado de la respuesta.

EJEMPLOS DE CODIFICACIÓN DE PREGUNTAS CERRADAS

Podría indicar, aproximadamente, ¿cuántos cigarrillos fuma Ud. al día?

Sólo 1 ó 2	1
Menos de medio paquete	2
Un paquete diario	3
Alrededor de paquete y medio	4
Dos paquetes de cigarrillos o más	5

¿Cómo calificaría Ud. su situación económica?

Pésima	1
Mala	2
Regular	3
Buena	4
Muy buena	5

Los *códigos* suelen figurar a la derecha de cada opción de respuesta.

En la investigación mediante *encuesta,* los *cuestionarios* están, en su mayoría, integrados por *preguntas cerradas*. Ello se debe a las grandes ventajas que proporciona este formato de pregunta (enfocado a la *estandarización* de la recogida de información). Entre las *ventajas* principales de las *preguntas cerradas* destacan las siguientes:

a) La rapidez y la comodidad de su registro. Resulta bastante más sencillo y rápido anotar la respuesta de una *pregunta cerrada* que la correspondiente a una *abierta*.
b) La posibilidad de centrar las respuestas de los encuestados a aquellas opciones consideradas relevantes y relacionadas con la cuestión que se pregunta.
c) La inmediatez de la grabación de las respuestas en el ordenador, una vez concluido el *trabajo de campo*.
d) Permite una mayor comparación de las respuestas, al encontrarse éstas expresadas en los mismos términos. Esto también ayuda a eliminar la vaguedad o ambigüedad de las respuestas.

No obstante, han de considerarse, igualmente, los graves *inconvenientes* de las *preguntas cerradas*. Éstos se agrupan en dos esenciales:

a) Las *preguntas cerradas* coartan las opciones de respuesta. Estas opciones, sin embargo, no siempre se ajustan a la variedad de respuestas posibles. Lo que revierte (negativamente) en la simplificación de la información obtenida.
b) El diseño de una *pregunta cerrada* es bastante más laborioso que el de una *pregunta abierta*. Exige, del investigador, el previo conocimiento de la realidad que investiga, su delimitación y medición expresa. No sólo ha de decidir cómo formular la pregunta, sino también, qué categorías de respuesta considerar, en qué nivel de *medición* (nominal, ordinal, de intervalo, de razón), y qué *códigos* asignar a cada respuesta.

A estos inconvenientes se suma la duda de si las distintas opciones de respuesta son igualmente interpretadas por todos los encuestados. La *estandarización* de las palabras, que se consigue igualando la literalidad de las preguntas y las respuestas, no implica necesariamente *estandarización* de los significados. A una misma respuesta los encuestados pueden atribuir significados diferentes. Respuestas como, por *ejemplo,* "mucho", "poco", "bastante", "viejo" o "joven", suelen provocar interpretaciones dispares de una persona a otra. Mientras que algunos encuestados pueden considerar que fumar 5 cigarrillos al día es "poco", otros, en cambio, pueden afirmar que es "mucho". Del ingenio y agudeza del investigador depende la reducción, o eliminación, de estas limitaciones de las *preguntas cerradas*.

• *Preguntas abiertas*

Como su nombre denota, las *preguntas abiertas* son aquellas que no circunscriben las respuestas a alternativas predeterminadas. Por lo que la persona puede expresarse con sus palabras.

EJEMPLOS DE PREGUNTA ABIERTA

¿Qué opina Ud. de las campañas electorales? ....
....
....
¿A dónde le gustaría ir de vacaciones? ....
....

Esta "libertad" concedida al encuestado lleva consigo un inconveniente importante: el coste en tiempo (y dinero) que supone la traducción de las respuestas "libres" en categorías que las resuman. Lo que se conoce como el *cierre de las preguntas abiertas.*

Una vez que los *cuestionarios* se han rellenado, ha de extraerse una *muestra* representativa de éstos para, a continuación, proceder a la *codificación* de las *respuestas abiertas*. Esta *muestra* puede oscilar entre el 20% y el 50% de los *cuestionarios* completados (Bourque y Clark, 1994).

De los *cuestionarios* extraídos al azar, una o varias personas encargadas deben, en primer lugar, transcribir, literalmente, las distintas respuestas emitidas. En segundo lugar, buscarán (en las diversas respuestas) términos comunes para, en función de ellos, agruparlas en un número reducido de *categorías*. El número de *categorías* resultante dependerá de:

a) La variabilidad de las respuestas.
b) Los objetivos de la investigación. Si se estima de interés una mayor especificación de las respuestas o, por el contrario, se busca su síntesis en un número reducido de *categorías* genéricas.

Cada *categoría* debe incluir un número considerable de respuestas similares. El contenido de éstas dictará el nombre (o *etiqueta)* que se dará a la *categoría*. A su vez, las *categorías* han de cumplir los requisitos comunes a la *codificación de preguntas cerradas* (exhaustividad, exclusividad y precisión). A cada una de ellas se le asignará, igualmente, un *código numérico,* que facilitará su tratamiento informático.

A la laboriosidad que supone la codificación de preguntas *abiertas,* se suma otro *inconveniente* importante: es más probable que se cometan errores en el registro y traducción de la información. El entrevistador (cuando el cuestionario no es autoadministrado) ha de anotar, literalmente, la respuesta del encuestado. Es decir, no debe introducir ninguna modificación que pudiera alterar su significado. Esto, obviamente, consume más tiempo de entrevista que señalar una *respuesta cerrada*. Asimismo, los codificadores deben procurar proporcionar *etiquetas* que se ajusten al significado común de las respuestas agrupadas.

Pese a estos inconvenientes, las preguntas *abiertas* se consideran de gran *utilidad:*

a) En *estudios exploratorios,* cuando no se dispone de un conocimiento previo suficiente del tema que se investiga.
b) Cuando el investigador no prevé todas las posibles respuestas a una determinada cuestión; o ésta precise de la enumeración de un listado extenso de respuestas.

EJEMPLOS DE PREGUNTA ABIERTA

¿Qué le gustaría a Ud. hacer ahora? ....
....
¿Qué programas de televisión ha visto Ud. este fin de semana? ....
....

c) Cuando se desea una mayor especificación de una respuesta dada con anterioridad como, por *ejemplo,* conocer los motivos de una determinada conducta.

EJEMPLOS DE PREGUNTA ABIERTA COMPLEMENTANDO UNA CERRADA

¿Tiene pensado cambiar de vivienda en el futuro?

Sí 1
No 2

→ ¿Por qué? ....

¿Ha participado Ud. en la organización de alguna de las actividades del centro?

Sí 1
No 2

→ ¿En cuáles? ....

d) Si se quiere conocer el valor numérico exacto de una determinada variable. La variable quedará medida a nivel de *intervalo.* Ello da opción a una mayor variedad de análisis estadísticos.

EJEMPLOS DE PREGUNTA ABIERTA NUMÉRICA

¿Cuántos hermanos tiene Ud.? ....
Años cumplidos ....
¿Cuántos años lleva casado? ....

En general, las *preguntas abiertas* presentan cuatro *ventajas* principales:

a) Proporcionan una mayor información (más específica y precisa) de las cuestiones que se investigan.
b) Su formulación resulta más sencilla.
c) Suelen ocupar menos espacio en el *cuestionario*.
d) Ofrecen al encuestado la posibilidad de expresarse en sus propias palabras.

En suma, la elección entre un formato de pregunta u otro dependerá de la combinación de tres factores básicos:

a) El *tiempo* y los *recursos* que el investigador quiera destinar a la *codificación* de *preguntas abiertas*.
b) El grado de *exactitud* que desee en las respuestas.
c) Su *conocimiento previo* del tema que investiga.

En el Cuadro 7.5 se resumen algunas de las ventajas y de los inconvenientes principales de cada formato de pregunta. El investigador deberá valorarlos, antes de optar por una u otra opción de pregunta-respuesta.

CUADRO 7.5. Ventajas e inconvenientes de las preguntas abiertas y cerradas.

	Tipo de pregunta	
	Abierta	*Cerrada*
Ventajas	Proporciona una información más amplia y exacta, expresada en los propios términos del encuestado. Fácil de formular. Suele precisar de menos espacio en el cuestionario.	Fácil de responder y de codificar. Reduce la ambigüedad de las respuestas. Favorece la comparabilidad de las respuestas. Requiere menos esfuerzo por parte del encuestado.
Inconvenientes	La codificación es más compleja y laboriosa. Más expuesta a errores en el registro de las respuestas y en su codificación. Su contestación exige más tiempo y esfuerzo por parte del encuestado (y del entrevistador, en su caso).	Su redacción exige un mayor esfuerzo y conocimiento del tema por el investigador. Limita las respuestas a opciones (o categorías) previamente acotadas. Las respuestas pueden tener diversas interpretaciones en los encuestados.

7.2.2. La formulación de preguntas

El éxito de una *encuesta* descansa bastante en su adecuación al proceso de medición; en cómo se hayan operacionalizado los *conceptos teóricos* en las preguntas concretas del *cuestionario*. De ahí la reiterada recomendación de no escatimar ni tiempo, ni esfuerzo, en la realización de las tareas preliminares de la *encuesta:* desde la formulación teórica del problema, a su traducción en objetivos específicos de investigación (que delimiten la información que se precisa reunir). Cuanto más precisos y claros estén los *objetivos,* más fácil será decidir las preguntas que conviene realizar.

Fowler (1988) recomienda que antes de diseñar un *cuestionario:*

a) Se ponga por escrito lo que la *encuesta* pretende alcanzar.
b) Se elabore un listado que incluya las *variables* a medir para alcanzar los *objetivos* que se pretenden.
c) Se bosqueje el *plan de análisis* a desarrollar. Para tal fin, el investigador ha de tener claro qué *variables* serán las *dependientes,* cuáles las *independientes,* y cuáles las de *control* (o intervinientes), en las distintas *hipótesis* a contrastar.

A ello habría que añadir la conveniencia de concretar la *población* a analizar, y cómo se administrará el *cuestionario* (mediante entrevista personal, telefónica o por correo). Ambos aspectos inciden también en la *formulación de las preguntas* de un *cuestionario*.

Además de la *revisión bibliográfica* precisa en cualquier proceso investigador, en la investigación mediante *encuesta* adquiere gran relieve la *indagación exploratoria,* previa al *diseño del cuestionario*. La consulta de expertos, de archivos de datos de encuesta, y de otras fuentes bibliográficas, resulta exigida, en busca de información que documente los aspectos más relevantes a cubrir en la investigación y el modo de hacerlo.

También es de utilidad la realización de algún *grupo de discusión,* o de varias *entrevistas en profundidad,* a miembros de la población que se pretende analizar. Especialmente, cuando el investigador apenas dispone de información específica sobre el *problema* de estudio y la *población* de interés (su vocabulario y percepción de la problemática que se investiga).

> "Desafortunadamente, la mayoría de nosotros probablemente rehusemos esta fase preliminar y nos disponemos directamente a escribir nuevas preguntas y a tomar prestadas otras de la literatura de encuesta" (Converse y Presser, 1994: 132).

Sudman y Bradburn (1987) recomiendan que, antes de crear nuevas preguntas, el investigador busque (en archivos de datos y materiales publicados) preguntas desarrolladas sobre la misma temática por otros investigadores. Si bien, es improbable que el investigador encuentre un estudio que cubra todos sus *objetivos.* Por lo que, su *cuestionario* podrá incluir algunas preguntas tomadas de una o varias encuestas precedentes, junto a

preguntas creadas originariamente por él, y otras que constituyan modificaciones o adaptaciones de preguntas formuladas por otros.

• Si el investigador opta por tomar *preguntas existentes,* deberá pretestarlas, por dos razones principales:

a) El lenguaje cambia constantemente.
b) El significado de las preguntas puede estar afectado por el contexto de las preguntas adyacentes en la entrevista (Converse y Presser, 1994: 133).

Una vez pretestadas, el investigador tendrá que decidir si tomarlas tal y como figuran enunciadas o, por el contrario, modificarlas. Si se inclina por su inalteración, podrá comparar sus resultados con los alcanzados en indagaciones anteriores. En concreto, podrá –de acuerdo con varios autores (Sudman y Bradburn, 1987; Bourque y Clark, 1994)–:

a) Replicar los hallazgos de un estudio en otra población o en una fecha posterior, y comparar los resultados.
b) Estimar la *fiabilidad* de las respuestas, en estudios realizados con poblaciones y contextos similares (donde no existan razones para esperar cambios).
c) Analizar la *tendencia,* para períodos de tiempo más largos, o donde se esperen cambios.

• Si opta, en cambio, por la *elaboración de nuevas preguntas,* existen una serie de criterios comúnmente aceptados por la mayoría de los autores. Estos criterios, o *"recomendaciones",* cabe resumirlos en los siguientes:

a) *Formular preguntas relevantes a la investigación.* Antes de redactar una pregunta, hay que valorar su utilidad para la consecución de los *objetivos* propuestos.

   Sudman y Bradburn (1987) recomiendan que, como regla general, el investigador se pregunte (cada vez que piense en una posible pregunta) *"¿por qué estoy preguntando esta pregunta?".*

   Siempre tiene que haber algún vínculo de la pregunta con el *problema de investigación.* En caso contrario, deberá descartarse la pregunta.
b) *Preguntas breves y fáciles de comprender, por las personas a las que van dirigidas.*

   "Sin duda, la mejor estrategia es utilizar preguntas breves, cuando sea posible, y que los entrevistadores den a los entrevistados tiempo suficiente para responder a las preguntas" (Converse y Presser, 1994: 93).

   A ello se suma la necesidad de conocer el nivel educativo y el vocabulario de la *población* a encuestar. La elección de las palabras que componen la

pregunta ha de supeditarse a las peculiaridades de la *población*. Por lo que, han de evitarse expresiones que puedan inducir a error.

Como regla, deben utilizarse palabras que sean comprensibles por los miembros de menor nivel educativo de la *muestra* elegida. Las frases también han de ser breves y sencillas. Las preguntas largas y complejas aumentan la probabilidad de que el encuestado se pierda y no las siga (Orenstein y Phillips, 1978).

c) *Evitar palabras ambiguas* (que carezcan de un significado uniforme). Esta recomendación adquiere mayor relevancia cuando el *cuestionario* se administra sin la presencia de un entrevistador, que aclare el significado de la pregunta y las respuestas.

Palabras habituales en el argot de los sociólogos (como "interacción social", "alienación", "socialización") no son plenamente comprendidas por la generalidad de los individuos (Newell, 1993).

Igualmente, términos como "trabajador", "mayor", "joven", "progresista", "mucho", "barato", "usualmente", pueden variar de acepción, dependiendo de a quién se pregunte. Lo que dificulta la interpretación de las respuestas.

d) *No emplear palabras que comporten una reacción estereotipada*. Si, por *ejemplo*, se desea conocer el grado de racismo de la población española, *preguntas directas* como "¿Podría Ud. decirme si es racista?", no aportarían la información buscada. En la sociedad actual, términos como "racista", "homosexual", "drogadicto" o "fascista", tienen una connotación peyorativa. Hecho que revierte en la baja proporción de personas que, abiertamente, reconocen que lo son.

Para obtener dicha información podría recurrirse a *preguntas indirectas,* relacionadas con el tema que se investiga. Por *ejemplo*, "¿Llevaría Ud. a su hijo a un centro escolar donde estudien niños gitanos?". Si bien, algunos autores desaconsejan el uso de *palabras hipotéticas,* como "¿Qué haría si...?", o "¿Le gustaría...?" .

Newell (1993) observa que lo que el encuestado dice que podría hacer cuando se enfrenta a una situación dada, no siempre expresa su conducta futura real. Existen preguntas que inevitablemente producen respuestas favorables (como, por ejemplo, "¿Le gustaría tener mayores ingresos?") por su deseabilidad social; y otras, en cambio, respuestas desfavorables. No obstante, el autor reconoce que la utilidad de las *preguntas hipotéticas* lo dicta el tema de estudio.

e) *Proporcionar respuestas flexibles, o atenuar la gravedad de la pregunta, cuando se aborden cuestiones que inhiban para transmitir una información veraz.* Hay que procurar que la pregunta no incomode al encuestado. Preguntas como "¿Alguna vez ha robado Ud. algo de un gran almacén?", suelen provocar rechazo en el encuestado. Sería mejor preguntar: "¿Alguno de sus amigos acostumbra a extraer productos en grandes almacenes?". Después, se formu-

larían *preguntas indirectas,* que indiquen si la persona suele, igualmente, practicar dicha actividad.

También puede optarse por solicitar al encuestado una *respuesta aproximada.* Preguntas sobre los *ingresos,* por *ejemplo,* suelen generar reticencia en los entrevistados a declarar la cantidad exacta. En estos casos, se recomienda la redacción flexible de la pregunta: "¿Podría indicar, aproximadamente, cuál es la cuantía de sus ingresos mensuales?".

Sudman y Bradburn (1987: 75-79) aconsejan (como estrategia para aumentar la probabilidad de informar *conductas no deseables),* "cargar deliberadamente la pregunta". La técnica a utilizar sería alguna de las siguientes:

1) *Todo el mundo lo hace.* Introducir la pregunta indicando que la conducta es muy corriente, con el propósito de reducir la amenaza de su revelación. Por *ejemplo,* "Incluso los padres más tranquilos alguna vez se enfadan con sus hijos. En los últimos siete días, ¿han hecho sus hijos algo que le enfadara?".
2) *Asumir la pregunta, y preguntar por su frecuencia y otros detalles.* Por *ejemplo,* "¿Cuántos cigarrillos fuma Ud. al día?". Si bien, los autores reconocen una desventaja importante en este tipo de pregunta: los encuestados que no realicen las actividades que se le preguntan podrían incomodarse, ante el supuesto de que sí las hagan. Esto repercutiría negativamente en su cooperación posterior.
3) *Uso de la autoridad para justificar la conducta.* Las personas pueden reaccionar de forma más favorable a una afirmación, si se atribuye a alguien que les gusta o que respetan. Por *ejemplo,* "Muchos médicos ahora afirman que beber vino reduce la probabilidad de sufrir un infarto, y favorece la digestión. ¿Bebe Ud. vino durante las comidas? ¿Con qué frecuencia?".
4) *Razones de por qué no.* Si a los encuestados se les da razones para no realizar conductas socialmente deseables (como votar, ponerse el cinturón de seguridad), se reduce su predisposición a no informarlas. Por *ejemplo,* "Muchos conductores afirman que llevar puesto el cinturón de seguridad es incómodo y dificulta la aproximación a los mandos del coche. Pensando en la última vez que Ud. se montó en su coche, ¿se puso el cinturón de seguridad?".
5) *Escoger marcos de tiempo apropiados.* Para conductas socialmente no deseables, los autores recomiendan comenzar con una pregunta como, por *ejemplo,* "¿Alguna vez ha cogido Ud. algo de una tienda sin permiso?".

   En cambio, si la pregunta es socialmente deseable, la estrategia inversa obtiene mejores resultados. A la persona le incomodaría admitir que no realiza conductas "deseables", como leer el periódico o ponerse el cinturón de seguridad. En estas situaciones, la pregunta correspondiente sería, por *ejemplo,* "Pensando en la última vez que Ud. subió en un coche, ¿se puso el cinturón de seguridad?". Esta pregunta se adecuaría más que la genérica "¿Alguna vez se ha puesto Ud. el cinturón de seguridad?".

f) *Formular la pregunta de forma objetiva (neutra), con objeto de no influir en la respuesta.* Deben evitarse *preguntas tendenciosas,* o que inciten a un tipo de respuesta. Preguntas como, por *ejemplo,* "¿No llevaría Ud. a su madre a una residencia de ancianos?"; además de incitar a dar una respuesta negativa, no desvelaría la actitud real de la población hacia sus mayores.

Asimismo, se recomienda que fórmulas como "La mayoría de las personas opinan que...", o la referencia a la opinión expresa de una autoridad (como "La Iglesia opina que..."), no antecedan una pregunta (salvo que se opte, deliberadamente, por estas fórmulas con el propósito de aumentar la probabilidad de informar sobre conductas no deseables, como se ha expuesto anteriormente). Estas introducciones de preguntas sesgan las respuestas, y hay que procurar evitar la inclusión de *sesgos* en las preguntas.

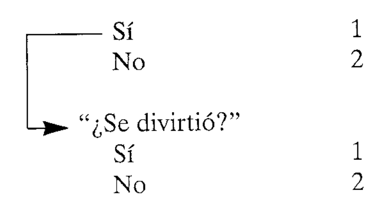
> "El investigador debe formular las preguntas para dar igual énfasis a todas las alternativas; para legitimar todas las variedades de opinión; para hacer que el encuestado estime que cualquier respuesta que escoja sea tan aceptable como cualquier otra" (Orenstein y Phillips, 1978: 219).

g) *No redactar preguntas en forma negativa.* La formulación negativa de una pregunta suele comprenderse peor que la formulación positiva. De manera especial, cuando se pide al individuo que manifieste su grado de acuerdo o desacuerdo ante determinadas cuestiones.

En vez de afirmar, por *ejemplo,* "No debería castigarse a los estudiantes que suspenden", podría preguntarse "A los estudiantes que suspenden ¿debería castigárseles?". De esta forma quedaría más claro el significado de una respuesta negativa.

h) *Las preguntas no deben referirse a varias cuestiones al mismo tiempo* (el principio de la idea única). Preguntas como "¿Cree Ud. que la sociedad actual es egoísta y competidora?", no facilitan la interpretación correcta de la respuesta. La persona no puede separar la doble mención de la pregunta. Lo mismo cabría decir de la pregunta "En las últimas navidades, ¿cenó con sus familiares y se divirtió?". Puede que la respuesta a la primera parte de la pregunta sea afirmativa, pero la correspondiente a la segunda, negativa. La pregunta comprende dos cuestiones distintas. Por tanto, habría que formular dos preguntas diferentes. Por *ejemplo:*

"En las últimas navidades, ¿cenó Ud. con sus familiares?"



Lo mismo cabe decir de las opciones de respuesta de una *pregunta cerrada*. Estas tampoco deberían contener dos o más ideas afines en una misma categoría.

i) *Evitar preguntas que obliguen a realizar cálculos mentales o a recurrir, con frecuencia, a la memoria.*

Este tipo de preguntas ponen en juego la *fiabilidad* de las respuestas e, incluso, la posibilidad de que el sujeto las responda.

En general, las preguntas que requieren el ejercicio de la memoria son más difíciles de responder, de una forma precisa. Sobre todo, cuanto más inusual o trivial sea el acontecimiento que se le pregunta.

Converse y Presser (1994: 102-104) sugieren algunas *"técnicas para aumentar la validez de informar sobre el pasado":*

1) Preguntar por acontecimientos que han sucedido en los últimos seis meses. Después puede incluso remontarse "más allá".
2) Estrechar el período de referencia al pasado más inmediato (como la última semana o ayer). En vez de preguntar, por *ejemplo,* "Realiza regularmente algún ejercicio físico?". Si se responde afirmativamente, "¿Cuántas horas a la semana?"; sería mejor preguntar: "¿Hizo ayer Ud. algún tipo de ejercicio físico". Si responde afirmativamente, "¿Qué tiempo dedicó?".
3) *Promediando.* Preguntas de "promedio" (por término medio), o concernientes a un día "típico", por *ejemplo,* suelen resultar más útiles que cuestiones relativas a un día concreto.
4) Tomar como referencia acontecimientos o fechas importantes del calendario, para datar acontecimientos de la vida personal. Por *ejemplo,* "Desde Año Nuevo, ¿ha recibido Ud. algún regalo?".

Sudman y Bradburn (1987) recomiendan, además, el uso de *"procedimientos de ayuda al recuerdo"*. Consisten en proporcionar, al encuestado, una o más señales de memoria, como parte de la pregunta. Estas "señales" pueden consistir en proponer algunos ejemplos en la pregunta, como por *ejemplo:* "¿A qué organizaciones pertenece Ud.? (por ejemplo, religiosas, sindicales, gubernamentales, etc.)"; o en mostrarle una *tarjeta* que contenga un listado de respuestas. Así, en vez de preguntar "¿Qué hace Ud. para relajarse?", podría mostrarse una *tarjeta,* que comprendiese distintas actividades de ocio y de deporte, y preguntarle si las practica o no:

	Sí	No
Ir al cine	1	2
Cenar en un restaurante	1	2
Ir de compras	1	2
Dar un paseo	1	2
Montar en bicicleta	1	2
Ver la televisión	1	2
Escuchar música	1	2

En caso de acudir a *"procedimientos de ayuda",* estos autores (Sudman y Bradburn, 1987: 37-43) aconsejan tomar algunas *precauciones:*

1) La *lista* de recuerdos que se proporcione debe ser tan *exhaustiva* como sea posible.
2) *Formular preguntas específicas.* El uso de cuestiones específicas ayudará a reducir la diferencias en la interpretación que puedan hacer los sujetos.
3) *Marcar un período de tiempo correcto.* Los períodos de dos semanas a un mes se adecúan más a los acontecimientos de escasa notoriedad. En cambio, para acontecimientos de mayor relieve (o notoriedad intermedia), los períodos de uno a tres meses resultan más apropiados.

j) *Redactar preguntas de forma personal y directa.* En lugar de preguntar, por *ejemplo,* "¿Qué sentiría Ud. si pasease por una barriada de chabolas?", se obtendría mayor información con la pregunta: "¿Ha paseado Ud., alguna vez, por una barriada de chabolas?"; si es así, "¿Qué sintió?".

Igualmente, cuando se analicen conductas que precisen una especificación temporal o numérica, se aconseja proporcionar categorías de respuesta específicas. Las expresiones "con frecuencia" o "regularmente" son demasiado vagas y ambiguas. Por lo que deberían evitarse. En cambio, las opciones de respuesta concretas (como "diariamente", "2-3 veces a la semana", "una vez a la semana", "dos veces al mes") resultan más adecuadas si, por ejemplo, se desea conocer la frecuencia de lectura de periódicos (Newell, 1993).

Las *preguntas específicas* suelen proporcionar una información más precisa que las *preguntas genéricas.* Estas últimas pueden, por el contrario, provocar una mayor variedad de interpretaciones. Razón por la que algunos autores (Converse y Presser, 1994) restringen su uso a las circunstancias en que interese la obtención de una medida global.

k) *Rotar el orden de lectura de las alternativas de respuesta* (cuando se prevea que su disposición pudiera afectar la respuesta). Esta recomendación adquiere mayor relevancia en la *encuesta telefónica.* Bosch y Torrente (1993) señalan que la *encuesta telefónica* es, particularmente, sensible al efecto *recency* (carácter reciente o novedad). Dicho efecto favorece, sobre todo, a las opciones de respuesta ubicadas en un determinado lugar (por ejemplo, al principio, o al final, de la lista de las categorías de respuesta).

• Las preguntas de un cuestionario pueden hacer referencia tanto a hechos objetivos, como a opiniones subjetivas y actitudes. Cuando el investigador desea analizar la *actitud* concreta de una persona, lo normal es que recurra a algún *procedimiento escalar* especializado en la *medición de actitudes* (más que a la formulación de varias preguntas independientes). Estos procedimientos le permitirán conocer no sólo la dirección de la *actitud,* sino también su intensidad.

Desde la publicación en 1928 del famoso y provocativo artículo de Thurstone en la revista *American Journal of Sociology* ("Attitude can be measured"), han ido apareciendo numerosas *escalas de actitudes*. Éstas consisten en una serie de afirmaciones que el investigador formula a los encuestados, para que estos indiquen su grado de conformidad (o de acuerdo) con las mismas.

Si el investigador está interesado en medir alguna *actitud,* al diseñar el *cuestionario* tendrá que elegir entre alguno de los *procedimientos escalares* existentes. De ellos, a continuación se comentan cuatro principales: la escala Thurstone, el escalograma de Guttman, la escala Likert y el diferencial semántico de Osgood.

a) Escala Thurstone

Contiene un conjunto de proposiciones relativas a una determinada *actitud*, expresadas en forma categórica (como aseveraciones). Al encuestado se le pide que indique su *acuerdo* o *desacuerdo* con cada proposición. El *promedio* de las respuestas resumirá su *actitud* ante el *problema* que se investiga.

Para su realización, es preciso que las proposiciones o *(ítems)* cubran el continuo de la *actitud.* Por lo que, el investigador deberá enunciarlas de manera que se correspondan con *distintas intensidades de la actitud* que miden.

EJEMPLO DE ESCALA THURSTONE

La medición de las "aspiraciones de los padres hacia el logro educativo de sus hijos" formaría una *escala Thurstone,* si se enunciasen proposiciones a modo de las propuestas por De Vellis (1991). Como, por ejemplo:

- Lograr éxito es sólo una forma de que mis hijos compensen mis esfuerzos como padre. — De acuerdo.................... En desacuerdo..............
- Ir a un buen colegio y obtener un buen trabajo son importantes, pero no esenciales en la felicidad de mis hijos. — De acuerdo.................... En desacuerdo..............
- La felicidad nada tiene que ver con lograr metas educativas o materiales. — De acuerdo.................... En desacuerdo..............

La aplicación de esta modalidad escalar es, no obstante, inusual en la práctica de la investigación social. Fundamentalmente, debido a su laboriosidad, pues precisa que previamente se hayan reunido un elevado número de *ítems* (de 100 a 150, según López, 1981). Posteriormente, un grupo de *jueces* (o especialistas) tiene que reducir la serie de *ítems* originales (a 20 ó 30), y asignarles un *valor escalar,* en función del grado de *ac-*

titud que representen. A este inconveniente principal se añade otra salvedad: la valoración de los *jueces* puede no ser coincidente con la *población* a la que se aplicará la escala.

b) El escalograma de Guttman

A diferencia de la modalidad escalar anterior, en ésta se reduce, sensiblemente, el universo de *ítems* (en torno a 30 enunciados, e incluso menos, si aparecen acompañados de otras preguntas en un *cuestionario),* y se elimina la *prueba de jueces* (basta con su previa comprobación en una *muestra* de la *población:* la *prueba piloto).* Los *ítems,* además, figuran ordenados de forma acumulativa y jerárquica. Por lo que, la afirmación de uno de ellos supone la corroboración de los precedentes.

A cada categoría de respuesta se le asigna una puntuación. La *puntuación* se fija de manera que los encuestados que hayan respondido favorablemente a una proposición ocupen una posición más elevada que aquellos que hayan respondido desfavorablemente.

Las categorías de respuesta pueden ser *dicotómicas* –"de acuerdo" (1), "en desacuerdo" (0); "sí" (1), "no" (0)–, o incluir más de dos opciones de respuesta (como en las *escalas Likert).* En este último caso, la *puntuación* más elevada se asigna al valor de la respuesta que sea más favorable a la *actitud.*

EJEMPLO DE ESCALOGRAMA DE GUTTMAN

Una versión de la escala de aspiración paterna a modo de *escalograma de Guttman* sería –siguiendo a De Vellis, 1991– la siguiente:

• Lograr el éxito escolar es la única forma de que mis hijos compensen mis esfuerzos como padre.	De acuerdo	1
	En desacuerdo	0
• Ir a un buen colegio y obtener un buen trabajo son importantes para la felicidad de mis hijos.	De acuerdo	1
	En desacuerdo	0
• La felicidad es más probable, si una persona ha logrado sus metas educativas y materiales.	De acuerdo	1
	En desacuerdo	0

Estas dos aproximaciones a la *medición de actitudes* (mediante afirmaciones *de acuerdo-en desacuerdo)* provocan en el encuestado, una mayor tendencia a mostrar *acuerdo,* indistintamente de cuál sea el contenido del *ítem.* Así, por ejemplo, Converse y Presser (1994) observan que esta incidencia es superior entre las personas de menor nivel educativo. Sea por esta u otras razones, las escalas tipo Thurstone y Guttman son poco habituales como formatos de preguntas en un *cuestionario.*

c) Escala Likert

Constituye uno de los formatos *escalares* más utilizados, cuando se desea preguntar varias cuestiones que comparten las mismas opciones de respuesta. En estos casos, se confecciona una *matriz de ítems* (o aseveraciones). A los encuestados se les pide que respondan a cada afirmación, escogiendo la categoría de respuesta que más represente su opinión.

A diferencia de las modalidades anteriores, en las *escalas Likert* normalmente existen cinco categorías para cada *ítem* ("muy de acuerdo", "de acuerdo", "indeciso", "en desacuerdo", "muy en desacuerdo"). Las *puntuaciones* (1, 2, 3, 4, 5, ó, a la inversa, 5, 4, 3, 2, 1) se asignan en conformidad con el significado de la respuesta para la actitud que miden. La *puntuación global* de la *escala* suele obtenerse sumando todas las puntuaciones registradas (incluidas las inversas).

EJEMPLOS DE PUNTUACIÓN DE CATEGORÍAS EN UNA ESCALA LIKERT

- Debería prohibirse fumar en todos los lugares públicos.

1	2	3	4	5
Muy de acuerdo	De acuerdo	Indiferente	En desacuerdo	Muy en desacuerdo

- El precio del tabaco debería reducirse.

5	4	3	2	1
Muy de acuerdo	De acuerdo	Indiferente	En desacuerdo	Muy en desacuerdo

Un problema importante de la *escala Likert* es que la misma *puntuación global* puede obtenerse de distintas combinaciones de respuestas (Procter, 1993). También, hay que tener presente que los procedimientos desarrollados por Thurstone, Guttman y Likert son métodos directos de *medición de actitudes* que se enfrentan a un problema común: el individuo puede descubrir la *actitud* que se mide y modificar su respuesta (si quiere ajustarse a las demandas del investigador o a las que perciba como socialmente deseables).

Al igual que el *escalograma de Guttman,* la *escala Likert* no precisa de la *prueba de jueces*. Basta con administrarla a una pequeña *muestra* de sujetos (unas 100 personas).

EJEMPLO DE ESCALA LIKERT

Como ilustración de una *escala Likert,* se extracta una pregunta del *cuestionario* aplicado en la investigación de Torres *et al.* (1994).

"A continuación le voy a leer una serie de frases. Me gustaría que me dijera (para cada una de ellas) si está muy de acuerdo, de acuerdo, indiferente, en desacuerdo, o muy en desacuerdo."

	Muy de acuerdo	*De acuerdo*	*Indiferente*	*En desacuerdo*	*Muy en desacuerdo*
• Preferiría no haber tenido hijos.	1	2	3	4	5
• Los hijos han sido una gran satisfacción en mi vida.	1	2	3	4	5
• Los hijos vienen con un pan bajo el brazo.	1	2	3	4	5
• Los hijos son un problema que nunca acabas de llevar.	1	2	3	4	5
• Por más que uno se esfuerce, al final los hijos salen como quieren.	1	2	3	4	5

Las proposiciones de este ejemplo miden actitudes diferentes hacia los hijos. Por esta razón, la *puntuación* de cada proposición deberá ser consonante con la *actitud* que mide, aunque se haya decidido que en el *cuestionario* conste la misma *puntuación* en cada *ítem.* Después, en la *grabación* de los *cuestionarios* contestados, se modificarán las *puntuaciones* según la intensidad de la *actitud* expresada.

d) El diferencial semántico de Osgood

En este último procedimiento escalar, se pide a los encuestados que marquen una de las posiciones (generalmente siete) que median entre dos *adjetivos polares.* Cada posición representa una opción de respuesta. De esta manera se miden los *sentimientos* (positivos o negativos) de las personas hacia un objeto concreto de *actitud.*

Para su elaboración se recomienda que los pares de adjetivos se alternen, de forma aleatoria, para que las respuestas positivas (o negativas) no caigan en el mismo extremo.

La *puntuación* se obtiene asignando "1" a la posición que indica la respuesta más negativa, y el "7", a la respuesta más positiva. La *puntuación global* será, igualmente, el promedio de todas las respuestas para la totalidad de los enunciados propuestos.

EJEMPLO DE DIFERENCIAL SEMÁNTICO DE OSGOOD

"En su opinión, ¿cómo definiría las amistades de sus hijos? Coloque una X en una de las siete posiciones comprendidas entre cada par de adjetivos polares." *(Mostrar tarjeta.)*

Estudiosos	- - - - - - -	No estudiosos
Desobedientes	- - - - - - -	Obedientes
Tímidos	- - - - - - -	Descarados
Callejeros	- - - - - - -	Caseros

Esta última modalidad de formato escalar resulta útil en situaciones donde la gente es probable que presente fuertes reacciones emocionales hacia una determinada cuestión (Henerson *et al.,* 1987).

7.2.3. La disposición de las preguntas en el cuestionario y su codificación

El *cuestionario* ha de diseñarse de forma que parezca *atractivo* y *cómodo* de responder. De manera especial, cuando éste sea *autoadministrado*. A tal fin, son varios los aspectos a cuidar. Éstos tienen que ver, sobre todo, con el formato del cuestionario y la secuencia de las preguntas en él.

• Respecto al *formato del cuestionario,* Sudman y Bradburn (1987) dan las siguientes *recomendaciones:*

a) *Utilizar el formato de libro*. Este tipo de formato facilita la lectura del *cuestionario* y la vuelta de página. Pero, también, previene la pérdida de páginas, y parece más profesional y fácil de seguir.
b) *Espaciar las preguntas*. La disposición de las preguntas sin apenas espacio entre ellas, con la intención de que el *cuestionario* parezca más corto, provoca efectos no deseados: una menor cooperación y mayor probabilidad de *errores* en su cumplimentación.
c) *En las preguntas abiertas, proporcionar espacio suficiente* para que se anoten las respuestas. El espacio dado suele interpretarse, en general, como indicador de la cantidad de información que se desea.

d) *Imprimir el cuestionario en color blanco o pastel.* También suele ayudar la utilización de *papeles de varios colores* para diferenciar las distintas partes del *cuestionario.*

En cuanto a la *impresión,* Newell (1993) aconseja, además, escoger una *impresión atractiva* para el documento, y que ésta sea fácil de leer. Si el presupuesto lo permite, utilizar papel de buena calidad.

e) *Numerar las preguntas.* Las preguntas principales suelen numerarse, consecutivamente, con números arábigos (de *1* hasta *n).* Las preguntas que figuran bajo un mismo encabezamiento se identifican, normalmente, por letras (A, B, C...).

f) *Evitar que las preguntas queden partidas entre páginas.* Una pregunta (incluyendo todas sus categorías de respuesta) nunca debería figurar entre dos páginas.

g) *Una pregunta larga,* que incluye varias partes, *no debería seguirse de una pregunta breve al final de la página.* Dicha pregunta suele, con frecuencia, omitirse por error.

h) *Proporcionar instrucciones al entrevistador,* en lugares apropiados a lo largo del *cuestionario.* Estas *instrucciones* han de poderse identificar con facilidad, mediante una impresión diferenciada del resto (acudiendo, por ejemplo, a letras mayúsculas, en cursiva, o a otros recursos tipográficos).

Las *instrucciones* se disponen delante de la pregunta, si tienen que ver con la manera de formular o responder la pregunta. Por el contrario, las instrucciones figuran detrás de la pregunta, si se refieren al modo de registrar las respuestas, o señalan al entrevistador cómo debería sondearlas.

También ayuda la preparación de *especificaciones* que acompañen al *cuestionario:* comentarios explicativos, y clarificaciones, en caso de complicaciones que puedan presentarse en el *trabajo de campo* (Babbie, 1992).

i) *Disponer las respuestas en sentido vertical.*

j) *En la entrevista personal, recurrir a tarjetas* (en *preguntas cerradas complejas)* para que el entrevistado pueda visualizar las distintas opciones de respuesta.

Cuando se empleen *tarjetas,* los entrevistadores deberán leer las preguntas, y las opciones de respuesta, en voz alta (aunque los encuestados estén visualizando la *tarjeta).* Esta precaución se toma por si algún entrevistado tiene problemas de lectura o de visión.

k) *En preguntas filtro, dar instrucciones expresas* (literales o mediante flechas), que apunten a la siguiente pregunta. Es importante que la instrucción figure inmediatamente después de la respuesta. De este modo es menos probable que el entrevistador (o el encuestado) la pase por alto.

Una *pregunta filtro* es aquella que se formula con anterioridad a otra (u otras), con la finalidad de eliminar a los sujetos a los que no procede hacer la pregunta siguiente.

EJEMPLOS DE PREGUNTA FILTRO

P.1. ¿Ha hecho amistades entre las personas que acuden al centro?

Sí 1
No 2

→ P.1.A. De ellas, ¿cuántos son varones y cuántas mujeres?

Varones ....
Mujeres ....

P.1. ¿Está Ud. jubilado?

Sí 1
No 2

Sólo si está jubilado
P.1.A. ¿A qué edad se jubiló?
....

Sólo si se jubiló antes de los sesenta y cinco años
P.1.B. ¿Cuál fue el motivo de su jubilación? ....
....
....

l) En los cuestionarios que sean administrados por entrevistadores, dejar espacio para que el entrevistador anote la duración de la entrevista, y cualquier incidencia que estime de interés.

m) *Preparar el cuestionario para el procesamiento de los datos.* Ahorra tiempo y dinero. Esta preparación del cuestionario supone dos actividades previas:

1) *La precodificación de las preguntas cerradas.* Véase lo señalado al respecto en el subapartado 7.2.1.
2) *Precolumnar el cuestionario entero;* es decir, asignar a cada pregunta o *ítem* una localización de *columna,* para su procesamiento informático.

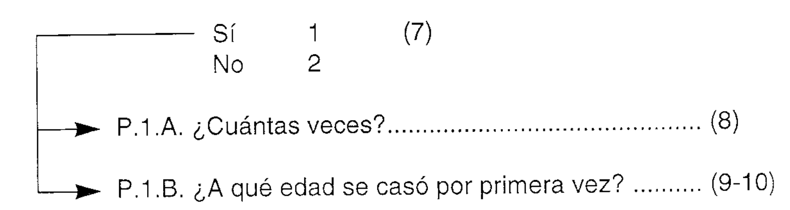
La *precolumnación* consiste en asignar a cada pregunta del *cuestionario* una o más *columnas,* dependiendo del número de respuestas posibles. Si la pregunta es *cerrada,* ya se conoce, con certeza, el número de columnas que requerirá. Pero, si la pregunta es *abierta,* debe de preverse el tipo y el número de respuestas posibles. A menos que la respuesta sea numérica, generalmente no se precisa más de dos columnas. Si son menos de diez las opciones de respuesta posibles, se asigna a la pregunta una sola columna. Si son diez o más las respuestas, dos columnas. En este último caso se codificarían las respuestas desde 00, 01, 02... hasta 99.

Cuando el encuestado puede dar más de una respuesta en una misma pregunta *(pregunta múltiple),* el número de columnas depende del número de alternativas que se pidan, y del tratamiento informático que se dé a esta columnación especial.

EJEMPLOS DE PRECOLUMNACIÓN DE PREGUNTAS DE UN CUESTIONARIO

- N.° de cuestionario .... (1-4)
- Comunidad autónoma .... (5-6)

P.1. ¿Ha estado Ud. casado?



P.2. De la siguiente lista de cualidades, por favor escoja las tres que más le gustan en una persona (pregunta múltiple). Mostrar tarjeta A.

	1.° (11-12)	2.° (13-14)	3.° (15-16)
Lealtad	01	01	01
Sinceridad	02	02	02
Tolerancia	03	03	03
Amabilidad	04	04	04
Simpatía	05	05	05
Humildad	06	06	06
Sencillez	07	07	07
Perseverancia	08	08	08
Entrega	09	09	09
Afabilidad	10	10	10

Los *códigos* correspondientes a las *columnas* suelen imprimirse en el margen derecho de la pregunta. Las primeras columnas habitualmente se dejan para los *datos de control* del *cuestionario* (número de estudio, número de cuestionario, localidad, duración de la entrevista, entrevistador, ...). Esto permite la identificación de cada *registro* (entrevistado) con cada *cuestionario,* una vez que la información se transforma en códigos numéricos para su posterior grabación en un *fichero de datos*.

Pero, téngase presente que no todos los *cuestionarios* son previamente *precolumnados*. Depende del *sistema de grabación* que se utilice. En los sistemas CADAC, por ejemplo, no se precisa de la *precolumnación*. Tampoco resulta

necesaria en las entrevistas no asistidas por ordenador, si a la hora de *grabar* los datos se recurre a programas como el DBASE, o se hace uso de la utilidad de DATA ENTRY en las versiones últimas del SPSS, por ejemplo.

n) *Siempre termina la entrevista con un "gracias"*. "Esto sería automático para la mayoría de los entrevistadores, pero es mejor que cada cuestionario termine con un gracias impreso" (Sudman y Bradburn, 1987: 259).

Estas mismas recomendaciones (salvo, quizás, la *precolumnación)* se hacen extensibles a la administración del *cuestionario mediante ordenador*. House (1985) las resume en tres esenciales:

a) Utilización de procedimientos estándar para indicar preguntas e instrucciones.
b) Emplear la misma forma para indicar partes diferentes de la pantalla (por *ejemplo,* exponer algunas instrucciones breves al entrevistador; debajo, las preguntas y las categorías de respuesta; y, a la derecha de las categorías, información adicional, si se dispone).
c) No llenar la pantalla con demasiado texto. Las líneas con información importante deberían estar separadas por líneas en blanco para, así, facilitar su visualización.

Por último, si el *cuestionario* es *autoadministrado,* su *formato* adquiere mayor relevancia. A lo expuesto, se añaden sugerencias sobre la *cubierta* y la parte posterior de la *cubierta* del *cuestionario* (Bosch y Torrente, 1993). Concretamente:

a) La *cubierta* ha de contener: el título del estudio, una ilustración gráfica, el nombre y la dirección de la entidad responsable.
b) La parte posterior de la *cubierta* comprende: una invitación a realizar comentarios adicionales, con un espacio en blanco destinado a su anotación; el agradecimiento por la colaboración en el estudio; y, si se piensa enviar copias de los resultados de la investigación, instrucciones de cómo solicitarlos.

El *orden de las preguntas* también constituye una parte relevante del *diseño de un cuestionario*. La calidad de las respuestas puede verse afectada no sólo por la *redacción* de las preguntas, sino también por el *contexto* inmediato en el que figure cada pregunta (entre qué preguntas se halla comprendida), y su *ubicación* en el *cuestionario* (al principio, en el medio o al final).

Existen unas *convenciones,* generalmente aceptadas, sobre el *orden o disposición de las preguntas en un cuestionario*. Éstas pueden resumirse en las siguientes:

a) *Comenzar el cuestionario con un mensaje de presentación*. Es conveniente que la presentación incluya la identificación de la institución a cargo de la in-

vestigación. Pero, también, una breve explicación de: por qué se realiza la *encuesta,* cómo se ha seleccionado al encuestado, qué tipo de información se precisa, el interés del estudio, y los beneficios de su participación.

A ello se suma la garantía del completo *anonimato* de las opiniones que en él se manifiesten. Se solicita la cooperación del encuestado, y se le agradece su colaboración de antemano.

Cuando el *cuestionario* se administra mediante entrevista, es el entrevistador el encargado de realizar esta introducción de forma verbal. En la *encuesta por correo,* esta introducción figura en forma de *carta de presentación,* que precede al *cuestionario.* En ella se detallan, brevemente, los aspectos referidos. No se aconseja que su extensión supere una página. Asimismo, es conveniente que la *carta* incluya la fecha en que debe remitirse el *cuestionario.* Generalmente. dos semanas después de la fecha de su envío. Puede incluirse, igualmente, un número de teléfono para cualquier aclaración que el encuestado precise.

A la *carta de presentación* se añaden las *instrucciones* a seguir para la cumplimentación del *cuestionario.*

b) En el *cuestionario* debe también reservarse un espacio para los *datos de control* (al principio y/o al final del mismo).

Esto es especialmente importante, cuando el *cuestionario* se administra *mediante entrevista* (telefónica o personal). Los *datos de control* incluyen: el número de orden del *cuestionario;* si hubo sustitución del entrevistado; duración de la entrevista; nombre, dirección y teléfono del encuestado; nombre del entrevistador; fecha y hora de la entrevista; y cualquier otro dato que se estime relevante para el *control* del *trabajo de campo.*

c) *Las preguntas iniciales deben despertar el interés del encuestado y ser, al mismo tiempo, sencillas de responder.* Es preferible comenzar con preguntas fáciles de responder, que resulten interesantes al encuestado, y que no provoquen en él ningún retraimiento. Después, poco a poco, se irán introduciendo las más complejas *(técnica del embudo).* A veces, no obstante, se invierte la secuencia: el *cuestionario* comienza con preguntas específicas, pasando gradualmente a cuestiones más generales *(embudo invertido).* Se elige esta última disposición de las preguntas cuando el investigador quiere asegurarse que el encuestado ha considerado determinados aspectos al dar su respuesta.

En las *encuestas mediante entrevista* puede ser de gran utilidad comenzar el *cuestionario* con *preguntas abiertas,* bastante genéricas, que traten con el tema principal del estudio. Esto puede ayudar a romper el hielo. Por el contrario, en los *cuestionarios autoadministrados* se desaconseja el uso de *preguntas abiertas.* Lo mejor es comenzar con *preguntas cerradas* sencillas. Como señalan Sudman y Bradburn (1987: 218):

> "Las preguntas abiertas, que requieren escribir más de unas cuantas palabras, se perciben como difíciles, como potencialmente embarazosas, debido a la posibilidad de cometer errores gramaticales."

En los *cuestionarios autoadministrados* se desaconseja, igualmente, seguir la *técnica del embudo*. Ésta pierde su eficacia, porque el encuestado puede leer todo el *cuestionario* antes de comenzar a responder las preguntas. Pero esto no limita la aplicación de procedimientos de *embudo invertido:* comenzar por preguntas específicas para, posteriormente, pasar a las genéricas.

d) *Las preguntas "claves"* (aquellas que se consideren de especial relevancia para la investigación) *deberían ubicarse en el centro del cuestionario*. Estas preguntas suelen situarse, de forma estratégica, en el tercio medio del *cuestionario*. Después de haber despertado el interés del encuestado, y haber generado un ambiente distendido, favorable a la aplicación del *cuestionario* (sobre todo cuando éste se administra mediante entrevista); pero, antes de que el cansancio del encuestado comience a tener efecto.

e) *Distribuir secuencialmente las preguntas que formen una batería*. Por "batería" se entiende un conjunto de preguntas confeccionadas sobre una misma cuestión, que se complementan, al enfocar distintos aspectos del mismo tema.

Es conveniente que las preguntas que forman una *batería* figuren juntas en el *cuestionario,* formando una unidad. Si bien, en ocasiones, hay que optar por distanciarlas, si se quiere evitar que el encuestado trate de responderlas de forma coherente (por haber percibido la relación existente entre las preguntas anteriores y las posteriores). No obstante, esta dispersión tendrá que hacerse procurando no confundir (o molestar) al encuestado. Como observan Phillips y Orenstein (1978: 223):

> "Si una serie de preguntas sobre temas religiosos de repente es interrumpida por preguntas sobre renta, los encuestados no sólo pueden experimentar desorientación, sino que también pueden volverse suspicaces acerca de los propósitos del estudio."

f) *Las preguntas que se perciban como amenazantes* (para el encuestado) *es mejor situarlas al final del cuestionario,* cuando la reacción a ellas no pueda afectar a las respuestas de preguntas posteriores.

De ahí que la práctica común sea disponer las *preguntas de identificación* (edad, nivel de estudios, ocupación, religiosidad, nivel de ingresos) al final del *cuestionario*. Ello se debe, sobre todo, a la reacción negativa que puede provocar en el encuestado proporcionar *datos de identificación* personal. De manera especial, los concernientes a sus ingresos. Lo que podría repercutir, negativamente, en su participación en la *encuesta*.

Esta repercusión se elimina situando las *preguntas de identificación* al final del *cuestionario*. Aunque, a veces, hay que colocar algunas de estas preguntas al principio del *cuestionario*. Por *ejemplo,* las preguntas acerca de las relaciones padres-hijos precisarán la obtención previa de algunas características sociodemográficas, como la situación familiar del encuestado (estado civil, si tiene o no hijos, u otras), que actúen de *filtro* de los encuestados que pueden aportar información.

La percepción negativa de estas *preguntas de identificación* (como curiosidad o intromisión) puede también atenuarse, si figuran introducidas por algún preámbulo que informe al encuestado de la necesidad de su respuesta para compararla con las respuestas de otros grupos de población.

Por último, conviene insistir en la conveniencia de *diseñar un cuestionario no muy extenso*. Una norma básica en la elaboración de un *cuestionario* es evitar fatigar al encuestado, para que no merme la calidad de sus respuestas. De ahí la reiterada recomendación de eliminar todas aquellas preguntas que se consideren repetitivas o no relevantes a los *objetivos* de la investigación.

No es el número de preguntas lo que determina la amplitud máxima de un *cuestionario*, sino la duración media de la entrevista. Por lo general, se aconseja que ésta no sobrepase la hora, y mejor aún si dura entre treinta y cuarenta y cinco minutos. No obstante, ha de matizarse que es el interés que despierte el tema en el encuestado (más que la longitud del *cuestionario* en sí), lo que repercute en su atención y sensación de cansancio.

> "Para temas notorios, las entrevistas personales pueden durar de una hora a hora y media, y son posibles cuestionarios por correo en torno a 16 páginas. Para temas no notorios, los cuestionarios por correo suelen limitarse de 2 a 4 páginas" (Sudman y Bradburn, 1987: 227).

De ahí que, cuando el investigador decida la *longitud del cuestionario,* deba fijarse, primordialmente, en la notoriedad o interés de las preguntas para el futuro encuestado.

7.2.4. La prueba o pretest del cuestionario

Como cualquier instrumento de *medición,* el *cuestionario* debe probarse antes de su aplicación definitiva. Para ello se escoge una pequeña *muestra* de individuos (normalmente inferior a 100 personas), de iguales características que la *población* del estudio.

El objetivo esencial es evaluar la adecuación del *cuestionario:* la formulación de las preguntas y su disposición conjunta. En concreto, se pretende comprobar que:

a) Las *preguntas* tienen sentido, se comprenden, y provocan las respuestas esperadas. El significado que el investigador da a una pregunta puede que no se corresponda con el significado que el encuestado percibe.

b) La *categorización* de las respuestas (de las *preguntas cerradas),* y su *codificación*, sea correcta. A veces, la *codificación* de las *preguntas abiertas* se efectúa después de la prueba o *pretest* del cuestionario (y antes de su aplicación definitiva), a partir de las respuestas obtenidas en dicho *pretest.*

c) La *disposición* conjunta del *cuestionario* (su secuencia lógica) sea adecuada, y que su *duración* no fatige al encuestado.
d) Las *instrucciones* que figuran en el *cuestionario* se entiendan, así como el formato de las preguntas *filtro*.

De los resultados del *pretest* se desprenderá la necesidad de revisar: algunas de las preguntas del cuestionario, la secuencia de temas, la inclusión de nuevas preguntas o la redacción de otras instrucciones que ayuden a su correcta aplicación. Como afirma Bowen (1973: 90):

> "Nadie puede escribir un buen cuestionario a menos que haya hecho algunas entrevistas. Sólo de esta forma puede darse cuenta de cómo pueden formularse las preguntas para que la gente las comprenda y cómo las preguntas que suenan sencillas no se entienden por la gente corriente."

El *pretest* también proporciona otro tipo de información de interés para el desarrollo de la *encuesta:*

a) El *porcentaje* aproximado de *"no respuesta"* que se obtendría en la *encuesta*. De especial relevancia en la *encuesta por correo*.
b) La *idoneidad* del *marco muestral* utilizado en la investigación.
c) La *variabilidad* de la *población,* respecto al tema que se estudia.
d) La *preparación* de los *entrevistadores* (en *encuestas personales* y *telefónicas).*
e) El *coste* aproximado del *trabajo de campo* (en tiempo y dinero).

7.3. El trabajo de campo en una encuesta

En el desarrollo de una *encuesta,* la fase del *trabajo de campo* (o de recogida de información), es la más delicada y costosa. Su coste aumenta en relación con la complejidad de la *encuesta* y la dimensión del *diseño muestral* elegido; de mayor relevancia en *encuestas personales*.

Salvo en la *encuesta por correo,* la administración de los *cuestionarios* exige la formación de la *red de campo,* integrada, en su mayoría, por el equipo de entrevistadores y de supervisores.

7.3.1. La formación de los entrevistadores

En las *encuestas mediante entrevista* (personal o telefónica), la calidad de la información recogida depende, en gran medida, de cómo los entrevistadores hayan cumplido su trabajo. Entre las *funciones del entrevistador* se encuentran las siguientes:

a) *Localizar* a los entrevistados.
b) *Motivarles* para conseguir su participación en la investigación. Es muy importante que el entrevistador inspire confianza en el entrevistado. Ello repercute en su participación y en la sinceridad de sus respuestas.
c) *Leer las preguntas* en su exacta formulación y en el orden en que aparecen en el *cuestionario;* es decir, tal y como fueron diseñadas.
d) *Comprobar* si la *respuesta* del entrevistado se adecua al objetivo de la pregunta. En caso afirmativo, deberá registrarla lo más exactamente posible. En caso negativo, tendrá que volver a formular la pregunta, y aclarar cualquier duda que pudiera tener el entrevistado.

> "El mayor peligro en la recogida de datos es que quien los recoja pueda 'dirigir' al encuestado, o registre los datos de una entrevista, informe u observación de forma selectiva. Esto puede evitarse con una buena preparación y supervisión de su trabajo" (Bourque y Clark, 1994: 34).

La *formación de los entrevistadores* ha de cubrir todos los aspectos implicados en la actuación del entrevistador. Desde la selección de los sujetos, hasta el registro de las respuestas. Además de información específica sobre preguntas concretas del *cuestionario*.

La duración del *período de formación* oscila entre dos y cinco días. Depende de la complejidad de la *encuesta,* de la formación previa y del tamaño del grupo de entrevistadores seleccionados. Períodos de preparación inferiores repercuten negativamente en la investigación. Fowler (1988: 117) afirma que "los entrevistadores formados durante menos de un día producen, significativamente, más errores de encuesta que aquellos que fueron preparados durante más tiempo".

Algunos estudios han mostrado también el alcance de algunas *características del entrevistador* (sexo, edad, étnia, nivel educativo y cultural) en la calidad de las respuestas obtenidas (véase Alvira y Martínez Ramos, 1985; Fowler, 1988; Bosch y Torrente, 1993). Si bien, su efecto se debilita dependiendo, fundamentalmente, de las características del estudio (qué temas trate) y de la formación del entrevistador.

7.3.2. La supervisión del trabajo de campo

En la *encuesta telefónica,* la *supervisión* de la entrevista es más inmediata y sencilla que en la *encuesta personal.* Las entrevistas se realizan desde un mismo edificio. Lo que facilita la consulta de cualquier incidencia del *trabajo de campo* con los supervisores (o personal responsable). Además, el uso de ordenadores para la recogida de información (sistema CATI) facilita el control inmediato de las respuestas y el registro de los *datos de control.* Todo lo cual revierte en una mejor *supervisión del trabajo de campo*.

También en la *encuesta personal* suele incluirse un espacio (en el cuestionario) para el registro de los *datos de control:* nombre del que responde, dirección, teléfono, localidad, si hubo o no sustitución, número de intentos de contactación, cómo se procedió a la selección del encuestado, duración de la entrevista. Esta información servirá al supervisor para comprobar la actuación del entrevistador. No sólo para comprobar que realmente se entrevistó a quien debía entrevistarse (y número de intentos habidos en su localización), sino también el tiempo de duración de la entrevista.

> "Los entrevistadores que completan las entrevistas demasiado rápido pueden no estar dando a los encuestados tiempo suficiente para oír las preguntas o considerar sus respuestas" (Bourque y Clark, 1994: 42).

Estos dos autores desconfian de las entrevistas realizadas con prontitud, y recomiendan perspicacia ante los datos que se muestren demasiado perfectos.

Las *funciones* principales *del supervisor* son las siguientes:

a) *La revisión de los cuestionarios.* El supervisor comprueba los datos de identificación, el porcentaje de respuestas, la codificación de las preguntas, la omisión de preguntas, la calidad de las respuestas, y el seguimiento de las *preguntas filtro,* fundamentalmente.
   Muchos de los *errores* en la cumplimentación del *cuestionario* pueden evitarse si el entrevistador, inmediatamente después de la entrevista, realiza una *revisión* completa del *cuestionario* recién concluido.
b) Conversar (si es factible) con el jefe de campo y los entrevistadores, para conocer cualquier incidencia habida en la administración del *cuestionario.*
c) Si se ha seguido un *muestreo por rutas aleatorias,* el supervisor examina las *rutas* complejas sobre el terreno. El objetivo es comprobar si realmente se entrevistó a quien tenía que entrevistarse.

A la vista de la información recabada, el supervisor realiza el *informe de supervisión.* En él muestra el material que ha examinado (número de *cuestionarios,* de *hojas de rutas,* y de *rutas* en el terreno), y cualquier aspecto que juzgue de interés en el desarrollo de la *encuesta.*

7.4. Ejemplo de un cuestionario aplicado en una investigación real

A modo de ilustración, se expone un *cuestionario* propio diseñado en 1992 para el *Estudio sociológico de los socios de centros municipales de la tercera edad. Análisis evaluativo del servicio: clientela, satisfacción de los usuarios y demandas latentes* (Cea y Valles, 1993). Esta investigación fue financiada por el Área de Servicios Sociales del Ayuntamiento de Madrid.

El estudio tenía como *objetivo* fundamental comprobar si los centros de la tercera edad cumplen el fin para el que han sido creados. Este objetivo genérico implicaba la consecución de cuatro *objetivos operativos:*

a) Analizar la demanda potencial y los niveles cuantitativos-cualitativos de cobertura de los Centros Municipales de la Tercera Edad (CMTE).
b) Describir el perfil del usuario típico.
c) Determinar el grado de satisfacción con el centro.
d) Detectar las demandas no cubiertas.

El planteamiento metodológico comprendía la utilización complementaria de tres instrumentos técnicos: la *entrevista abierta* (a responsables próximos a los centros), el *grupo de discusión* (con personas mayores socios y no socios de los centros), y la *encuesta* (a los socios de los CMTE).

El *cuestionario* se elaboró teniendo en cuenta la información recabada, previamente, a través de *entrevistas abiertas* con profesionales de los centros. El *cuestionario* diseñado se reproduce a continuación. El estudiante puede hacer el ejercicio de comprobar hasta qué punto las consideraciones teóricas expuestas en este capítulo se llevaron a la práctica en este estudio.

CUESTIONARIO ESTUDIO SOCIOS CENTROS MUNICIPALES
DE TERCERA EDAD

Área de Servicios Sociales y Comunitarios. Ayuntamiento de Madrid.
Carrera de San Francisco, 10. 28005 MADRID. Tel.: 588-32-82

Un equipo de estudiantes y profesores de la Universidad Complutense de Madrid estamos realizando una encuesta entre los socios de centros municipales de tercera edad, para el Área de Servicios Sociales del Ayuntamiento. Ya hemos estado en el centro "...", y nos hemos presentado a la asistente social responsable del centro ("..."), a la animadora ("...") y a la Junta Directiva ("..."). Ha sido usted seleccionado/a, entre otros cientos de personas, para que nos conteste a unas preguntas. No se preocupe, pues las preguntas son muy fáciles y nadie sabrá que usted las ha respondido. Con su colaboración y la de todas las personas entrevistadas se podrá hacer un estudio que servirá para mejorar los centros de mayores.

N.º cuestionario ☐☐☐ Distrito ....

Centro Municipal Tercera Edad ....

ANOTAR Sexo: Varón 1
Mujer 2

Año de nacimiento: 19 ....

P.1. Para empezar, ¿podría usted decirme si vive...?

Solo/a	1
Con su esposo/a	2
Con su esposo/a e hijos	3
— Con sus hijos	4
— Con otros familiares	5
Otra situación (especificar)	

P.1.A. ¿Vive usted siempre con el mismo hijo o familia) o sólo a temporadas?

Siempre	1
A temporadas	2

P.2. Este domicilio en el que reside Vd. habitualmente, es

Alquilado por Vd. o esposo/a	1
De su propiedad o de esposo/a	2
Alquilado por su hijo	3
De propiedad de su hijo	4
De un familiar	5
Otra situación	

P.3. ¿Cuántos metros cuadrados tiene la vivienda?

□□□

P.3.A. ¿Cuántos dormitorios tiene la vivienda?

□□

P.4. Dígame el número de personas que viven ahora en esta vivienda contándose usted.

□□

P.5. Desde que usted formó su propio hogar, ¿ha vivido en alguna otra vivienda, algún otro barrio o localidad? (Si han sido varias, la última)

— Siempre aquí	1
— En otra vivienda mismo barrio	2
— En otro barrio de Madrid	3
— En otra localidad de la Comunidad de Madrid	4
— En otra Comunidad Autónoma	5
— En el extranjero	6

P.5.A. ¿Desde qué año lleva viviendo en este domicilio?

Desde el año 19 ....

(A TODOS)

P.6. ¿Tiene pensado cambiar de vivienda en el futuro? (la familia en su conjunto).

Sí	1
No	2

P.6.A. ¿Por qué piensan cambiarse?

....

....

....

....

P.7. ¿Diría Vd. que se encuentra Muy, Bastante, Poco o Nada satisfecho con...?

	Muy	Bastante	Poco	Nada
Su vivienda	1	2	3	4
El barrio	1	2	3	4
Los vecinos	1	2	3	4

P.8. ¿Tiene Vd. en su vivienda...?

	Sí	No
Teléfono	1	2
Televisión	1	2
Radio	1	2
Lavadora	1	2
Frigorífico	1	2

P.9. Y, ¿tiene Vd. en su vivienda...?

	Sí	No
Agua caliente	1	2
Cuarto de baño completo	1	2
Aseo (lavabo y retrete)	1	2
Retrete sólo	1	2
Calefacción central/individual	1	2
Radiadores eléctricos	1	2
Estufas de butano	1	2
Brasero/estufa de carbón/madera	1	2
Aire acondicionado	1	2

P.10. ¿Tiene el edificio ascensor?

Sí	1
No	2

P.10.A. ¿En qué planta vive (Baja, 1.ª, 2.ª, etc)?

....

P.11. ¿Cómo calificaría Vd. el estado de su actual vivienda?

Buen estado................................ 1
Necesitaría alguna pequeña reparación (pintura, etc.)........... 2
Reparaciones mayores............... 3

P.12. Respecto a la relación con sus vecinos del edificio, ¿diría Vd. que se lleva bien con...?

Todos 1
La mayoría de ellos 2
Unos pocos solamente 3
Ninguno 4

P.12.A. ¿A qué se debe?...
....
....

P.13. Entre sus vecinos del edificio hay alguien con el que pueda Vd. contar en caso de que...

	Sí	No
Necesitase algún recado (compra pequeña)..................	1	2
Necesitase que alguien le acompañe al médico................	1	2
Tuviese una pequeña avería en la vivienda..........................	1	2
Enfermedad	1	2

P.14. Normalmente, ¿qué suele Vd. hacer durante la mañana? *(Anotar las tres actividades principales.)*

Entresemana	Domingos
1.........................	1..........................
....	....
2.........................	2..........................
....	....
3.........................	3..........................
....	....

P.15. ¿Dónde hace cada una de las comidas del día?

	Casa	Fuera Casa	Indicar dónde	No Toma
Desayuno	1	2	....	3
Comida	1	2	....	3
Merienda	1	2	....	3
Cena	1	2	....	3

P.16. Y por las tardes, ¿qué suele hacer, normalmente? ¿Y los domingos?

Entresemana	Domingos
1.........................	1..........................
....	....
2.........................	2..........................
....	....
3.........................	3..........................
....	....

P.17. ¿Cuánto tiempo hace que es Vd. socio/a del Centro Municipal de la Tercera Edad "..............."?

Desde el año 19 ....

P.18. ¿Cómo conoció este Centro? A través de.......

Vecinos... 1
Familiares 2
Amigos.. 3
Parroquia 4
El Ayuntamiento (JMD)........... 5
Otros (especificar) ....
....
....

P.19. Dígame ¿con qué frecuencia acude Vd. a este Centro?

A diario o casi 1
Tres o cuatro días a la semana ... 2
Uno o dos días a la semana 3
Nunca o casi nunca...................... 4

P.19.A. ¿Va Vd. por la mañana o por la tarde?

Sólo por la mañana..................... 1
Sólo por la tarde.......................... 2
Mañana y tarde 3

P.20. Por favor, ¿recuerda qué días de la semana pasada ha ido? (si L, M, X, J....)

L M X J V S D

(EN CASO DE NO HABER IDO NINGÚN DÍA)

P.20.A. ¿A qué se debió? ....
....
....

(EN CASO DE HABER IDO ALGÚN DÍA)

P.20.B. ¿Cuáles son las tres razones principales por las que acude?

1...
2...
3...

(A TODOS)

P.21. En este centro, ¿qué actividades hay?

....
....

P.22. De esas actividades, ¿cuáles realiza Vd.?

....
....

P.23. ¿Ha participado Vd. en la organización de alguna de las actividades del centro?

Sí	1
No	2

P.23.A. ¿Cuáles?

....
....

P.23.B. ¿Por qué no?

....
....
....

P.24. Los Centros Municipales de la Tercera Edad suelen contar con una serie de Servicios para sus socios. Dígame si conoce o no en su centro los que le cito a continuación. (En caso afirmativo) ¿Querría calificarlos de 1 a 10 según la opinión que tenga de ellos? (Si valora menos de 5) ¿Por qué?

	Conoce		Útil			
	Sí	No	Sí	No	Valora	Por qué
Podología......	1	2	1	2	....	....
Peluquería......	1	2	1	2	....	....
Comedor........	1	2	1	2	....	....
Cafetería	1	2	1	2	....	....
Gimnasio	1	2	1	2	....	....
Biblioteca.....	1	2	1	2	....	....
Tall. pintura...	1	2	1	2	....	....
Tall. cestería..	1	2	1	2	....	....
Tall. macramé	1	2	1	2	....	....
Prensa..........	1	2	1	2	....	....
Vídeo.........	1	2	1	2	....	....

P.25. ¿Hay algún otro servicio en su centro que NO le haya mencionado?

Sí	1
No	2

P.25.A. ¿Cuál/es?

	Utiliza			
	SÍ	NO	Valora	Por qué
....	1	2	....	....
....	1	2	....	....
....	1	2	....	....

P.26. De los servicios que NO tiene el Centro, ¿cuál/es echa Vd. más en falta?

....
....
....

P.27. Tengo aquí una lista de cosas que la gente suele hacer en un centro. Dígame, para cada una de ellas, si Vd. suele hacerlas y con qué frecuencia ("S" siempre que voy/hay, "V" a veces, "R" rara vez), (SÍ "V" ó "R") ¿Por qué?

	Sí	No	S	V	R	Por qué
Jugar cartas ..	1	2	1	2	3	....
Ver televisión	1	2	1	2	3	....
Desayunar	1	2	1	2	3	....
Comer	1	2	1	2	3	....
Tomar algo....	1	2	1	2	3	....
Ir a charlas.....	1	2	1	2	3	....
Baile	1	2	1	2	3	....

P.27.A. ¿A cuántas excursiones/salidas organizadas por el centro ha ido usted, en el último año?

Fuera de Madrid ....
Sin salir de Madrid municipio ....

P.28. ¿Hay alguna actividad que no organiza el Centro y que a Vd. le gustaría que se hiciera en el Centro?

Sí	1
No	2

P.28.A. ¿Cuáles?

....
....

P.29. En relación con alguna de las actividades del centro, ¿ha tenido Vd. dificultades de encontrar sitio (mesa/silla) o plaza para realizarlas?

Sí 1
No 2

P.29.A. ¿En qué actividades?

....

....

P.30. Teniendo en cuenta las actividades que Vd. realiza y los servicios que el centro le proporciona, ¿está Vd. muy ("M"), bastante ("B"), poco ("P") o nada ("N") satisfecho del centro?

	Muy	Bastante	Poco	Nada
Centro	1	2	3	4

P.31. ¿Recomendaría este Centro a un familiar o amigo?

Sí 1
No 2

P.31.A. ¿Por qué?

....

....

P.32. ¿Cuánto tiempo tarda en llegar al Centro desde su casa?

.... (minutos)

P.33. ¿Cómo va Vd. al Centro?

Andando	1
Autobús	2
Metro	3
Otro medio	4

P.34. Aparte de este centro, ¿hay algún otro centro de tercera edad o sociocultural que usted conozca?

Sí 1
No 2

P.34.A. ¿Cuáles?

P.34.A.	P.34.B.		P.34.C.			
Centros	Sí	NO	M	I	L	N
....	1	2	1	2	3	4
....	1	2	1	2	3	4
....	1	2	1	2	3	4

P.34.B. ¿Es socio?

....

P.34.C. ¿Asiste con mayor ("M"), igual ("I") o menor frecuencia ("L") o, por el contrario, no asiste ("N")?

(HABLANDO DE NUEVO DEL CENTRO MUNICIPAL "......")

P.35. ¿Pertenece Vd. a la Junta Directiva del centro?

Ahora no, pero antes sí	1
No, nunca he sido de la Junta.....	2
Sí, sí pertenezco	3

P.36. ¿Está Vd. **M**uy, **B**astante, **P**oco o **N**ada satisfecho con la labor de la actual Junta Directiva del Centro? ¿Y con la anterior?

	M	B	P	N	Por qué
Con la actual	1	2	3	4	....
Con la anterior	1	2	3	4	....

P.36.A. ¿Conoce a algún miembro de la Junta Directiva?

Sí 1
No 2

P.37. ¿Ha hecho amistades entre las personas que acuden al centro?

Sí 1
No 2

P.37.B. De ellas, ¿cuántas son hombres y cuántas mujeres?

Varones ....
Mujeres ....

P.38. ¿Cree Vd. que la Junta Municipal de Distrito dedica suficiente dinero a los Centros Municipales de la Tercera Edad?

Sí 1
No 2

P.39. ¿Es Vd. beneficiario de alguno de los programas del Ayuntamiento como.....

	SÍ	NO
Ayuda a domicilio	1	2

Otros (especificar) ....

....

P.40. Hablando de su familia, ¿tiene Vd. hijos, hermanos, sobrinos y/o nietos? ¿Dónde residen habitualmente?

	En casa		En Madrid		Fuera de Madrid	
	SÍ	NO	SÍ	NO	SÍ	NO
Hijos	1	2	1	2	1	2
Hermanos	1	2	1	2	1	2
Sobrinos	1	2	1	2	1	2
Nietos	1	2	1	2	1	2

P.41. (SÓLO SI TIENE HIJOS QUE NO CONVIVAN CON EL ENTREVISTADO, SEGUN LA P.40.) ¿Por qué medios suele Vd. contactar con ellos?, y ¿con qué frecuencia, de cada manera? (AL MENOS CON UNO DE LOS HIJOS, CON EL QUE MÁS SE TRATE.)

	D	S	M	A	N
Va a visitarles	1	2	3	4	5
Vienen a visitarme	1	2	3	4	5
Teléfono	1	2	3	4	5
Carta	1	2	3	4	5

("D" a diario o casi, "S" una vez a la semana o varias, "M" una vez al mes o varias, "A" una vez al año o varias al año, "N" con menos frecuencia o nunca".)

P.42. (SÓLO SI TIENE FAMILIARES QUE NO CONVIVAN CON EL ENTREVISTADO SEGUN LA P.40.) ¿Por qué medios suele Vd. contactar con ellos?, y ¿con qué frecuencia, de cada manera? (A LOS QUE MÁS VE.)

	D	S	M	A	N
Va a visitarles	1	2	3	4	5
Vienen a visitarme	1	2	3	4	5
Teléfono	1	2	3	4	5
Carta	1	2	3	4	5

P.43. ESTUDIOS DEL ENTREVISTADO:

Ninguno y no sabe leer ni escribir	1
Ninguno, pero sabe leer y escribir	2
Primarios	3
Secundarios	4
Estudios de 3^er grado, medios	5
Estudios de 3^er grado, superiores	6

P.44.¿Podría decirme cuál era su ocupación principal antes de jubilarse? (o en caso de no estar jubilado) ¿A qué se dedica? (especificar categoría laboral.)

....

....

P.45. ¿Cobra algún tipo de pensión?

Sí 1
No 2

(SI "Sí"), P.45.A. ¿De qué cuantía, al mes? ....

....

(A TODOS)

P.46. Algunas personas de su edad realizan pequeños trabajos, para ayudarse económicamente, ¿está usted en este caso?

Sí 1
No 2

P.46.A ¿Cuántas horas dedica a ese trabajo al día?

....

P.47. ¿Tiene usted algún tipo de enfermedad o dolencia?

Sí 1
No 2

P.47.A. ESPECIFICAR ....

....

....

P.48. ¿Conoce Vd. a alguna de las personas que trabajan en relación con la tercera edad, en su centro/distrito?

	SÍ	NO
Asistente social (responsable del programa 3ª edad)..................	1	2
Animador/a sociocultural	1	2

Otras (ESPECIFICAR)...

....

....

....

P.49. ¿Conoce Vd. los Servicios Sociales de su Junta Municipal?

SÍ 1
NO 2

P.50. ¿Estaría Vd. dispuesto a dedicar parte de su tiempo a colaborar, como voluntario, en alguna de las actividades que le cito?

	SÍ	NO
Asistente social (responsable del programa 3.ª edad)..............	1	2
Visitar ancianos	1	2
Atención a niños sin hogar	1	2
Reinserción de drogadictos	1	2
Alfabetización de adultos	1	2
Monitor de algún taller de manualidades en el centro...	1	2
Hacer pequeños arreglos/reparaciones en el centro	1	2

NOMBRE ENTREVISTADOR:
....

LUGAR DE LA ENTREVISTA:
....

DURACIÓN ENTREVISTA:
....(minutos)

NOMBRE Y TEL. DEL ENTREVISTADO:
....

OBSERVACIONES:
....
....

Lecturas complementarias

Bosch, J. L. y Torrente, D. (1993): *Encuestas telefónicas y por correo*. Cuaderno Metodológico n.º 9. Madrid, CIS.

Clemente, M. y Fernández, I. (1992): "La medición de las actitudes". En Clemente (comp.): *Psicología social: métodos y técnicas de investigación social*. Madrid, Eudema, pp. 302-323.

Dillman, D. (1978): *Mail and telephone surveys: the total design method*. New York, Willey.

Fowler, F. J. (1988): *Survey research methods*. Beverly Hills, Sage.

García Ferrando, M. (1994): "La encuesta". En García Ferrando *et al.* (comps.): *El análisis de la realidad social*. Madrid, Alianza, pp. 147-176.

Lavrakas, P. J. (1993): *Telephone survey methods: sampling, selection and survey*. Newbury Park, Sage.

Sudman, S. y Bradburn, N. M. (1987) *Asking questions*. San Francisco, Jossey-Bass.

Ejercicios Propuestos

1. ¿Qué se entiende por error de cobertura? ¿Qué modalidad de encuesta es más vulnerable a este tipo de error? Razone la respuesta.
2. El cuestionario que figura en el apartado 7.4 no está precodificado. Efectúe la precodificación y precolumnación del cuestionario entero. Además, compruebe si las consideraciones teóricas expuestas en este capítulo se han llevado a la práctica en el cuestionario.
3. En la colección *Estudios* del Ministerio de Asuntos Sociales, en los números 21 y 22 se adjunta el cuestionario íntegro de ambos estudios. Compárense los dos cuestionarios, destacando las diferencias entre ellos.
4. Elabore diez ítems que midan la actitud de los jóvenes hacia sus mayores. Especifique el procedimiento escalar seguido, y señale qué variaciones supondría la aplicación de otros procedimientos escalares.
5. ¿Qué cualidades buscaría en un entrevistador? Justifique la respuesta.
6. Diseñe un cuestionario sobre la "ocupación del tiempo libre". Describa sucintamente las actuaciones seguidas en su elaboración.

8

LA EXPERIMENTACIÓN COMO ESTRATEGIA DE INVESTIGACIÓN SOCIAL

La *experimentación* constituye una *estrategia de investigación* más habitual en el campo de la psicología social que en la sociología, propiamente. Su uso suele vincularse a *objetivos* de investigación *explicativos* y/o *evaluativos;* al análisis de relaciones causales, sobre todo. De hecho, éste es el único propósito que autores como Orenstein y Phillips (1978) o, más recientemente, Hakim (1994) le reconocen. Específicamente, el "establecer si un factor 'X' dado tiene algún efecto en otro factor 'Y', o si los cambios en una variable producen cambios en otra" (Hakim, 1994: 101).

No obstante, otros autores (Roberts, 1983), añaden una segunda *finalidad* fundamental a la *experimentación,* independiente de la primera: la comprobación de la efectividad de soluciones alternativas a problemas concretos. Por *ejemplo,* comprobar si un tratamiento resulta más eficaz que otro para reducir la agresividad. El cumplimiento de esta finalidad no demanda, necesariamente, la realización de un estudio teórico-empírico sobre las causas de la agresividad (aunque siempre se beneficiará de su realización previa, pues el conocimiento de las causas ayudará a una mejor *evaluación* de las soluciones). El *experimento,* en este caso, adquiere más un componente pragmático que teórico. En cambio, la aplicación de la *experimentación* en la consecución de la primera finalidad (la indagación de *relaciones causales)* se hace con fines eminentemente teóricos: la contrastación de hipótesis de teorías existentes, y la configuración de *modelos causales.*

Pero, ¿qué se entiende por experimentación?, ¿en qué consiste un experimento?, ¿qué variedades de experimentos existen?, ¿cuáles son las ventajas y los inconvenientes de la experimentación? A estos y otros interrogantes se tratará de dar respuesta en los apartados que dividen el presente capítulo.

8.1. Características esenciales de la experimentación

La *experimentación* puede definirse como un modo de hacer investigación, basado en el *control* e *intervención* del investigador en la "realidad" que analiza. Este *control* se dirige a la comprobación de los efectos de la variable *manipulada* por el investigador (la *variable independiente),* en la ocurrencia de la variable cuya variabilidad se investiga (la *variable dependiente).*

No se trata de una técnica de recogida de datos, sino de una *estrategia* específica de investigación. Al igual que en las otras *estrategias,* en la *experimentación* la información se obtiene mediante la aplicación de una o varias *técnicas de obtención de datos* (como la *observación sistemática,* el *cuestionario,* la *entrevista).* Si bien, hay que precisar que estas *técnicas* ahora se orientan hacia la medición de los cambios provocados por el "estímulo experimental" en la aparición del fenómeno que se analiza.

Sin duda, la *experimentación,* se muestra como una *estrategia* que se adecua al estudio de la *causalidad.* A ello contribuye el elevado *control* ejercido durante su realización, con el objeto de eliminar cualquier explicación alternativa a las *relaciones causales* observadas.

El estudio de la *causalidad* no sólo exige el cumplimiento de los tres criterios fundamentales del concepto de *causalidad* enunciados, en el siglo XVII, por Hume (en su *Tratado sobre la naturaleza humana).* A saber:

a) Comprobar la *contigüidad* entre la *causa* y el *efecto* (las dos variables implicadas han de covariar).
b) La *precedencia temporal* de la *causa* sobre el *efecto* (la variable independiente ha de acontecer antes que la dependiente).
c) La *conjunción* constante entre la *causa* y el *efecto* (siempre que se presenta la "causa" sucede el "efecto" y, a la inversa, cuando no aparece la "causa", tampoco se observa el "efecto").

También ha de cumplirse el criterio enunciado por John Stuart Mill en el siglo XIX: la exclusión de otras posibles alternativas potenciales (véase Díez Medrano, 1992). En la medida en que este último requisito se cumpla, la indagación efectuada muestra *validez interna* –como ya se expuso en el Capítulo 3–.

En la investigación mediante *encuesta*, por ejemplo, la eliminación de explicaciones alternativas a una *relación causal* se hace mediante el *control a posteriori* (es decir, una vez que la información se ha recogido), durante el proceso de análisis. Por el contrario, en la *experimentación,* el *control* de explicaciones alternativas se realiza *a priori* (antes de la recogida de información), cuando se diseña la investigación. Ello es posible por las propios componentes o *características* básicas que definen la *experimentación:*

a) La manipulación experimental.
b) El control del investigador.
c) La aleatorización.

• *La manipulación experimental*

El investigador crea una situación (artificial) que le permita comprobar la influencia causal de una o más variables *(independientes)* en otra (la variable *dependiente).* La variable *independiente* (cuyo efecto se trata de medir) adquiere la forma de *tratamiento* o *estímulo experimental.* El número de *tratamientos* varía en función de cuántos valores o categorías incluya la variable *independiente.*

A veces, al investigador sólo le interesa la comparación entre dos grupos (el *experimental* y el de *control).* Pero, en otras ocasiones, el *diseño experimental* se complica, al incluir más de una variable *independiente* en el análisis, o ésta comprende más de dos valores. De la diversidad de *diseños experimentales* tratará el apartado 8.2.

• *El control del investigador*

La *manipulación experimental* se halla condicionada a las posibilidades de controlar (o eliminar) la influencia de cualquier variable que pueda incidir en la relación causal entre dos variables (y que no interese al estudio). Sin este requisito, no cabe hablar de *influencia causal.* Como afirma Jones (1985: 286):

> "Cuando se está interesado en comprobar si una manipulación determinada en algún aspecto del entorno tiene algún efecto, se necesita controlar, o mantener constante, otros aspectos del entorno."

Claro que no es fácil *controlar* todos los factores que afectan a una *relación causal.* Pero el investigador deberá procurar controlar (o mantener constante) todos aquellos que, de acuerdo con sus indagaciones previas, o a su juicio, pudieran afectar.

• *La aleatorización*

La *aleatorización* es un componente esencial en todo *experimento.* Se halla presente en dos momentos fundamentales del proceso experimental: cuando se asigna cada sujeto a alguno de los grupos formados; y cuando se asigna cada grupo a una de las condiciones experimentales (Morales *et al.,* 1981).

En cualquier *experimento* habrá, al menos, dos *grupos de tratamiento:* el *grupo experimental* (al que se aplica el tratamiento, cuyo efecto trata de medirse), y el *grupo de control* (al que se le priva de dicho tratamiento, o se le asigna un tratamiento *placebo).* Mediante la formación de este segundo grupo, se intenta "controlar" el efecto de factores ajenos al estímulo, o *tratamiento experimental,* que puedan afectar a los resultados del estudio (características internas y externas a los individuos incluidos en la investigación). Para ello ha de procurarse la total *equivalencia* inicial de ambos grupos. Es decir, que los integrantes de cada uno de ellos compartan las mismas características. En

lo único en que deben diferir es en el nivel de *tratamiento* aplicado posteriormente. En la medida en que esto se consiga, podrá diferenciarse los efectos (en la variable *dependiente)* provocados por la manipulación de la variable *independiente*.

La *equivalencia* puede alcanzarse de dos formas principales:

a) Mediante la *igualación* razonada de los sujetos, en aquellas variables (o características conocidas) que pudieran afectar al valor de la variable *dependiente*. Se trataría, por tanto, de eliminar (o neutralizar) su influencia, garantizándose su presencia, y en la misma magnitud, en los distintos grupos diferenciados. De esta forma, aquellos sujetos que compartan dichas características serán asignados *aleatoriamente,* en igual proporción, a los distintos grupos. Por el contrario, aquellos que no las compartan serán eliminados del *experimento*.

   El problema principal de este procedimiento de *control,* de variables ajenas al estudio, es que el investigador nunca podrá estar plenamente seguro de si realmente ha "igualado" a los distintos grupos, en todas las variables que pudieran afectar a los resultados de la investigación.

b) La *aleatorización* (o asignación aleatoria de los sujetos a los grupos de tratamiento) destaca como otra opción que ayuda a la equivalencia inicial de los grupos.

   Al seguirse un procedimiento de asignación grupal aleatorio, se da a todos los individuos, que participan en el *experimento,* la misma probabilidad de pertenecer a cualquiera de los grupos. Esto favorece la *equitatividad* en el reparto de la *muestra* en los distintos grupos. Aunque pudieran existir leves diferencias entre ellos, la *aleatorización* garantiza la equivalencia "aproximada" de los grupos en un número elevado de características.

Si el investigador desea comprobar (antes de la aplicación del *tratamiento)* la equivalencia inicial de los grupos, puede acudir a un *cuestionario,* u otro instrumento, que le permita medir la presencia de iguales características en los grupos diferenciados.

Una vez concluido el *experimento,* el investigador procede a la *comparación* de los resultados en los distintos grupos, para así comprobar el efecto del *tratamiento* aplicado. La *comparación* de los resultados grupales en la variable *dependiente* puede realizarse mediante un análisis de *varianza* u otra técnica de análisis.

En la investigación mediante *experimento,* es frecuente el *análisis de la varianza,* porque permite la medición estadística de las diferencias de las *medias grupales* en la variable *dependiente*. Los *tests de significatividad* (el estadístico *t,* en el análisis de la *varianza simple;* y *F,* en el análisis *múltiple* de la varianza) indicarán si las diferencias observadas entre las *medias grupales* son o no significativas estadísticamente. Es decir, si éstas se deben o no a diferencias en la composición de los grupos.

Si los resultados del *experimento* muestran no ser estadísticamente significativos, no podrá afirmarse que el *tratamiento* ha tenido algún efecto. En cambio, cuando las

diferencias entre los grupos de tratamiento son estadísticamente significativas, implica que dichas diferencias no han de atribuirse a discrepancias (o divergencias) en la composición de los grupos. Por lo que, pueden imputarse a efectos distintos de la variable *independiente* en la *dependiente* (en consonancia con las *hipótesis* de investigación).

En todo ello influye no sólo el *tamaño de la muestra* analizada, sino también el procedimiento seguido en la distribución de los individuos en los grupos. Indudablemente, la presencia de la *aleatorización* en el desarrollo del *experimento* reducirá –como indican Saxe y Fine (1981)– la probabilidad de que los hallazgos entre las condiciones experimentales se deban, en realidad, a características particulares del entorno, del experimentador o de la población expuesta al *estímulo experimental*.

EJEMPLO DE EXPERIMENTO

Para ilustrar lo expuesto en este apartado, pensemos en un posible *experimento*. El *objetivo* de la investigación es comprobar la relación existente entre "pedagogía docente" y "formación del alumnado". En concreto se trataría de analizar si la manera como el profesor imparte la asignatura *(variable independiente)* influye en el nivel de conocimiento finalmente adquirido por el alumno *(variable dependiente)*.

A tal fin se disponen dos grupos de alumnos. Uno sería el *grupo experimental*, y el otro conformaría el *grupo de control*. Por lo que se tendría un *diseño experimental básico*. Los grupos estarían constituidos por alumnos lo más "iguales" posible. Quiere esto decir, que han de compartir las mismas características. La única distinción entre ellos se deberá a la "manipulación experimental": el profesor encargado de la docencia.

Al *grupo experimental* se le asignará un profesor practicante de la docencia "participativa" (en la que se conjuga teoría y práctica). Al *grupo de control*, en cambio, se le privará de las nuevas orientaciones pedagógicas, adjudicándole un profesor anclado en la práctica docente tradicional (limitada a las "clases magistrales"). Al final del experimento, los alumnos de ambos grupos deberán cumplimentar el mismo examen. Para ello, los dos profesores tendrán que impartir el mismo contenido del programa docente (a partir de la misma bibliografía). Los profesores sólo diferirán en la forma de exponer la materia, no en los contenidos impartidos.

Para que las diferencias en las calificaciones de los integrantes en ambos grupos puedan atribuirse al efecto de la variable manipulada, previamente habrá que controlar la incidencia de otras variables que puedan, a su vez, influir en la variable *dependiente*. Este control de *variables perturbadoras* se materializa en:

a) El control de características personales y sociales de los alumnos que participan en el experimento.
b) El control de características ambientales, o contextuales, que puedan influir en el desarrollo del experimento.

• *El control de las características de las unidades experimentales*

Con el propósito de que las diferencias entre los grupos no puedan atribuirse a características de sus integrantes, éstas deberán neutralizarse. En caso contrario, su influencia no

permitirá la comprobación del efecto concreto de la variable *independiente*, que centra el objetivo de la investigación.

Imaginemos que las calificaciones finales del *grupo experimental* superan las alcanzadas por el *grupo de control*. Si previamente no se ha efectuado ningún tipo de control, no podrá deducirse que los resultados correspondan a la manipulación realizada (la pedagogía docente). Puede que el *grupo experimental* se halle integrado por alumnos de mayor cociente de inteligencia, mejor formación académica, y mayor motivación al logro del éxito, que aquellos que forman el *grupo de control*. Tal vez sean alguna o varias de esas características diferenciales las que realmente influyan en los conocimientos finalmente adquiridos por el alumno durante el experimento. En caso de ser esta hipótesis cierta, seguiría habiendo diferencias entre los grupos, aunque en ambos se hubiera aplicado una misma pedagogía docente.

Para eliminar estas u otras *explicaciones alternativas*, el investigador deberá haber diseñado un experimento que garantice la *equiparación inicial de los dos grupos de tratamiento*. Como ya se ha expuesto, este "control" puede hacerse de dos formas básicas:

a) *Igualando* a los miembros de ambos grupos, de forma que el perfil de sus integrantes sea el mismo. A tal fin, el investigador elegirá, deliberada y razonadamente, las características que han de presentar las personas para participar en la investigación. Estas características se distribuirán, en igual proporción, en ambos grupos (por ejemplo, que en ellos exista un mismo porcentaje de hombres y de mujeres; de personas con distinto cociente de inteligencia, expediente académico y motivación al logro del éxito; y que puedan dedicar un mismo número de horas diarias al estudio). La asignación de los individuos, que compartan las mismas características, a los grupos se hará de forma *aleatoria*.

   Esta modalidad de control se enfrenta a una *limitación* importante: el investigador nunca podrá estar plenamente seguro de haber neutralizado la influencia de todas las posibles variables *perturbadoras*. Siempre cabrá la posibilidad de incidencia, en el proceso causal, de otras variables no controladas en la investigación. Como pueden ser, por ejemplo, algunas relativas a la situación familiar y económica de los alumnos (que les permita la dedicación exclusiva al estudio; la compra de libros y otros materiales escolares; el disponer de un espacio y de un ambiente favorecedor al estudio).

b) Si el investigador opta, en cambio, por la elección exclusiva de las unidades muestrales siguiendo procedimientos de *selección aleatorios* (por ejemplo, un muestreo aleatorio sistemático con las unidades muestrales potenciales), es más probable que alcance la "equivalencia" inicial (aproximada) de ambos grupos en un amplio número de variables. Ello se debe a que la *aleatorización* favorece la *equitatividad* en el reparto de la *muestra*.

Si el investigador desease comprobar, antes de comenzar el *tratamiento experimental* (que el profesor comience a impartir la asignatura), la igualdad inicial de los grupos, puede pasarles un *cuestionario* que le permita trazar la composición previa de los grupos.

• *El control de las características contextuales*

Aparte de la consideración de las características del alumnado, deberá asimismo neutralizarse la influencia de factores *contextuales*. Las clases serán impartidas, en ambos grupos, a

la misma hora del día, y con la misma duración. Las aulas serán exactamente iguales, con la misma capacidad (igual número de alumnos), y características de audición y luminosidad.

Dentro del edificio, la ubicación de las aulas será la misma, para evitar la influencia de su localización en la atención del alumno en la clase. Por ejemplo, que una de las aulas esté situada al exterior, con una mayor exposición a ruido ambiental que la otra aula; o que se halle orientada al oeste. Lo que, obviamente perjudicaría si la clase se impartiera después de comer, e hiciera una tarde soleada. Ello favorecería el sopor, tras el almuerzo, y la pérdida de atención del alumno a la exposición del profesor.

Estos y otros factores contextuales deberán controlarse. Cuantas más *explicaciones alternativas* puedan rechazarse, mayor seguridad habrá en la *influencia causal* de la *variable independiente* analizada.

Este minucioso *control "a priori"*, que exige la práctica de un *experimento*, contribuye a la *artificialidad* atribuida como característica definidora de la investigación experimental. Si bien, la existencia de *control* se convierte en requisito básico en cualquier comprobación de *hipótesis causales*. Lo que convierte a la *experimentación* en vehículo idóneo para el análisis de la *causalidad*.

8.2. Los diseños experimentales

Cuando se diseña un *experimento*, el investigador se enfrenta a la elección de una de las distintas variedades de *diseños experimentales*. Éstos, a partir de la aportación pionera y fundamental de Campbell y Stanley (1970), en la clasificación de los diseños de investigación, y de otros autores posteriores (véase Morales *et al.*, 1981), se dividen en la tipología esquematizada en el Cuadro 8.1.

CUADRO 8.1. Modalidades de diseños experimentales.

1. EXPERIMENTO INTERSUJETOS.
   - A. *De dos grupos aleatorios.*
     1. Diseño solamente después.
     2. Diseño antes y después.
     3. Diseño Solomon (combinación de los anteriores).
   - B. *De más de dos grupos.*
   - C. *Diseños factoriales.*
     1. Con dos variables independientes (2*2, 2*3, 3*3...).
     2. Con más de dos variables independientes (2*2*2, 3*3*2...).
2. EXPERIMENTO INTRASUJETOS.
   - A. *Dos tratamientos experimentales, un solo grupo.*
   - B. *Varios tratamientos experimentales, un solo grupo.*
   - C. *Diseños factoriales intrasujetos.*

8.2.1. El experimento intersujetos

Constituye la modalidad de *experimento* más habitual en la investigación social. Como su denominación expresa, supone la formación de distintos grupos de sujetos, a los que se aplican tratamientos diferentes, para comprobar sus efectos. Comprende las tres variedades siguientes:

a) Experimentos con dos grupos aleatorios.
b) De más de dos grupos.
c) Experimentos factoriales.

• *Experimentos con dos grupos aleatorios*

Representa la modalidad experimental más básica. Supone la formación aleatoria de sólo dos grupos. A cada uno de ellos se aplican tratamientos distintos, en consonancia con los valores o categorías diferenciadas en la variable *independiente:* únicamente dos categorías contrapuestas.

Un grupo sería el *grupo experimental* (aquél al que se da el *tratamiento* cuya influencia se desea comprobar) y, el otro, el de *control* (que sirve –como su nombre indica– para "controlar" los efectos de variables *perturbadoras).* Los sujetos se distribuyen, de forma aleatoria, en cada uno de ellos. A esta variedad de diseño corresponde el ejemplo anteriormente expuesto.

Dentro de este grupo de experimentos, se distinguen dos tipos principales, dependiendo del número de veces en que se ha medido la variable dependiente:

1) *Diseño solamente después.* Cuando la variable *dependiente* sólo se mide (en los dos grupos formados) después de la aplicación de los *tratamientos.* Al *grupo de control* puede, incluso, habérsele privado de tratamiento.

   El *inconveniente* fundamental de este tipo de diseño experimental es que, previamente, no se ha comprobado la equivalencia inicial de los dos grupos. Se confía en que ésta queda garantizada por haberse seguido un procedimiento de asignación aleatoria de las unidades muestrales a los grupos.
2) *Diseño antes y después.* Este diseño experimental es más complejo que el anterior. Implica la comprobación de la variable *dependiente* en ambos grupos (el *experimental* y el de *control),* tanto antes de la aplicación del *tratamiento* como después del mismo. De esta forma el investigador se asegura de que ambos grupos son equivalentes antes de aplicar los tratamientos. Lo que permite poder atribuir las diferencias observadas (en el valor de la variable *dependiente),* entre la primera y la segunda medición, al efecto de la variable *independiente.*

   Esta *premedición* del valor de la variable *dependiente* ayuda a una mejor comprobación del efecto del *tratamiento* aplicado, al conocerse el valor de la variable *dependiente* antes y después de la *intervención experimental.* Además,

puede conocerse la cantidad exacta de variación de una prueba a otra, en los dos grupos de tratamiento.

No obstante, esta *premedición* también lleva consigo un *inconveniente* importante: que los sujetos experimentales perciban la finalidad del estudio. Su conocimiento puede influir en su actuación posterior, no ajustándose a su proceder natural.

De la agudeza e ingenio del investigador dependerá el enmascarar el objetivo del estudio, con la finalidad de evitar el efecto no deseado que su conocimiento pudiera tener en la respuesta de los sujetos al *tratamiento*. Así, por *ejemplo*, si la premedición *(pretest)* se realiza mediante un *cuestionario*, habrá que incluir una amplia variedad de preguntas, que logren despistar al encuestado sobre el propósito del *cuestionario*.

3) *Diseño Solomon*. Este tipo de diseño es una combinación de las dos modalidades anteriores. Razón por la cual se incluye en este grupo de *diseños experimentales*.

La denominación de este diseño proviene del investigador que lo ideó. Consiste en la formación aleatoria, no de dos grupos, sino de cuatro: dos grupos *experimentales* y otros dos de *control*. A uno de los grupos *experimentales* y a otro de *control* se los mide antes y después de la aplicación del *tratamiento*. Por el contrario, a los otros dos grupos (uno *experimental* y otro de *control)* sólo se les mide el valor de la variable *dependiente*, cuando el *tratamiento* ha concluido.

Procediendo de este modo puede comprobarse el efecto que la *medición* inicial pudiera tener en los resultados de la investigación. A ello se suma otra *ventaja* importante: la posibilidad de comprobar de formas diferentes una misma *hipótesis* de investigación (Morales *et al.*, 1981).

- *Experimentos de más de dos grupos*

Cuando la variable *independiente* incluye más de dos valores, el *diseño experimental* será de más de dos grupos. El número de grupos dependerá del número de categorías o de valores diferenciados en la variable *independiente*.

Al igual que en los experimentos de dos grupos, en éstos también se trata de comprobar la relación causal existente entre una única variable *independiente* y una sola variable *dependiente*.

EJEMPLO DE EXPERIMENTO INTERSUJETOS DE MÁS DE DOS GRUPOS

Si en el ejemplo anterior, en vez de haber diferenciado dos tratamientos (pedagogía docente "participativa" y práctica docente "tradicional"), se hubieran distinguido tres (incluyen-

do una modalidad de práctica docente intermedia entre las dos referidas), el *diseño experimental sería de tres grupos.* Ello favorecería no sólo la medición del efecto de tratamientos contrapuestos, sino también la graduación de la influencia de distintos valores de la variable *independiente* en la *dependiente.*

• *Diseños factoriales*

Si al investigador le interesa analizar, al mismo tiempo, las relaciones causales existentes entre más de una variable *independiente* y una sola variable *dependiente,* su *diseño experimental* será *factorial.*

En los *diseños factoriales* se mide la influencia causal de *varias variables independientes,* cada una de ellas incluyendo un número variado de valores. El número de *grupos de tratamiento* estará en función del número de variables independientes consideradas, y de cuántos valores comprendan éstas. La razón se halla en que cada grupo se configura a partir de un valor de cada variable independiente. Por *ejemplo,* si se estudiase la influencia de dos variables independientes, y cada una de ellas incluyera dos valores, se tendría un *diseño factorial 2*2*; es decir, habría *cuatro grupos de tratamiento* diferentes. Si en vez de dos, fueran tres las variables independientes, y las dos primeras incluyeran dos valores, mientras que la tercera presentase tres valores, habría *2*2*3 = 12 grupos de tratamiento.*

Conforme aumenta el número de variables *independientes,* y los valores en ellas incluidos, mayor complejidad adquiere el *diseño factorial.* Obviamente, la *muestra* se distribuirá de forma aleatoria en los distintos grupos de tratamiento. Para ello se seguirán los mismos criterios de asignación a los tratamientos, a los que ya se ha hecho mención.

Estos diseños experimentales permiten la contrastación de varias *hipótesis causales* en una misma investigación. Los efectos de las variables *independientes* pueden, asimismo, comprobarse por separado y de forma conjunta. Lo que ayuda al análisis de efectos recíprocos de dos o más variables *independientes.*

EJEMPLO DE EXPERIMENTO INTERSUJETOS FACTORIAL

Se quiere estudiar el efecto que en la atención del alumno en clase (*variable dependiente)* tiene la manera cómo se imparte la asignatura (sólo teoría, sólo práctica, conjugando teoría y práctica) y la duración temporal de la clase (de dos horas, hora y media y una hora de duración). El *diseño experimental* sería *factorial,* con *9 grupos de tratamiento* formados aleatoriamente. Cada uno de los grupos resultaría de la combinación de dos de las categorías de las *dos variables independientes* analizadas. Los *grupos* serían los siguientes:

		FORMA DE IMPARTIR LA CLASE (B)		
		Sólo teoría	Teoría y práctica	Sólo práctica
Duración de la clase (A)	Dos horas	A_1B_1	A_1B_2	A_1B_3
	Hora y media	A_2B_1	A_2B_2	A_2B_3
	Una hora	A_3B_1	A_3B_2	A_3B_3

Al primer grupo, por ejemplo, se le asignaría una docencia exclusivamente teórica, siendo la duración de la clase de dos horas; al segundo grupo, la docencia es práctica y teórica, y la duración de la clase es, igualmente, de dos horas; y así con el resto de los grupos.

Esta combinación de tratamientos ayudará a la medición y graduación del *efecto conjunto* de las *dos variables independientes* consideradas.

8.2.2. El experimento intrasujetos

Este segundo grupo de *diseños experimentales* difiere del anterior en que los distintos tratamientos se aplican a *un único grupo* de sujetos. Se somete a las mismas personas, en momentos diferentes, a *varios tratamientos,* con la finalidad de analizar su reacción diferencial a los mismos. Las disparidades observadas en el valor de la variable *dependiente,* una vez concluido el *tratamiento,* se considera como una medida de su efecto.

EJEMPLO DE EXPERIMENTO INTRASUJETOS

En el ejemplo anterior, el *diseño experimental* sería *intrasujetos* si, para el mismo objeto (el análisis conjunto de la práctica docente y la duración de la clase), se observase a un único grupo, y no a nueve.

Las nueve modalidades de tratamiento referidas se aplicarían al mismo grupo de sujetos, espaciando un lapso de tiempo entre cada tratamiento.

Este tipo de diseño experimental comparte los *inconvenientes* comunes a cualquier análisis de una misma población de forma reiterada. Como los ya comentados (Ca-

pítulo 3) de *desgaste de la muestra,* con el consiguiente aumento de la "no respuesta": y los *sesgos* provocados por las mediciones anteriores en las posteriores (se da a la persona más oportunidades para adivinar el objetivo de la investigación, con la consiguiente modificación de su conducta; o que la persona quiera ser consonante con su actuación anterior, u otra razón).

Como resulta difícil –cuando se aplican *varios tratamientos* a los mismos sujetos– separar los efectos provocados por distintos tratamientos, el investigador debería controlar. especialmente, el orden de presentación de los tratamientos. Aquellos tratamientos que pudieran provocar más *sesgos* en las respuestas a tratamientos posteriores, deberían aplicarse al final del *experimento.* Los primeros tratamientos dados serían, precisamente. aquellos que provocasen menores *efectos reactivos* en la *muestra* observada.

Aparte de la graduación de los *tratamientos* (en función de su menor a mayor *reactividad),* el investigador podría también optar por *dividir la muestra,* en dos submuestras al menos. A cada una de ellas se le aplicarían los mismos tratamientos, pero en orden inverso. Esto permitiría medir los efectos reactivos de los distintos *tratamientos.*

Como en los *experimentos intersujetos,* en los *intrasujetos* se distinguen varias modalidades, dependiendo del número de variables independientes y de los valores o categorías que éstas incluyan. En función de ello, se diferencia entre:

a) *Dos tratamientos experimentales, un sólo grupo.* Sólo existe una variable *independiente,* y ésta comprende dos valores. A los sujetos de la *muestra* se les mide el valor de la variable *dependiente* antes y después de los dos tratamientos. Las diferencias que se observen entre ambas mediciones se atribuyen al efecto de los valores de la variable *independiente* en la *dependiente.* Por tanto, hay sólo un *grupo experimental,* y éste actúa, a su vez, de *grupo de control.*
b) *Varios tratamientos experimentales, un sólo grupo.* Ésta es la modalidad experimental, cuando la variable independiente incluye más de dos valores.
c) *Diseños factoriales intrasujetos.* Si hay más de una variable independiente, cuyos efectos conjuntos quieren medirse.

8.3. El control de explicaciones alternativas en la investigación experimental

Si por algo se distingue la *investigación causal* es por el riguroso *control,* que en ella se realiza, del proceso de investigación. Este *control* se dirige a la eliminación de cualquier explicación alternativa que invalide las conclusiones del estudio. Campbell y Stanley (1970) y, después Cook y Campbell (1979), fueron los autores que dieron forma al concepto de *validez interna* en referencia a este problema. Los distintos criterios de *validez* a considerar en cualquier *diseño de investigación* se expusieron en el capítulo 3, a cuya relectura remito.

A lo ya expuesto, aquí se añaden diversas *fuentes de invalidez* (indicadas por los autores citados) a controlar, si se pretende que el experimento presente *validez interna.* Las *fuentes de invalidez* principales son:

a) Maduración.
b) Historia.
c) Instrumentación.
d) Mortalidad experimental.
e) El efecto de la regresión.
f) Efectos de la medición inicial.

• *Maduración*

Con el paso del tiempo, las personas cambian en sus formas de pensar y de actuar. Este cambio o *maduración* de las personas puede incidir en los resultados del *experimento*. Sobre todo, si se analiza a la misma población durante un largo período de tiempo (diseños *intrasujetos).*

El investigador puede observar cambios entre una medición y otra (antes y después del *tratamiento)* y atribuirlos, erróneamente, al efecto del *tratamiento,* cuando en realidad se deben a la propia evolución del individuo.

El uso de un *grupo de control* ayuda al "control" de esta fuente de invalidez. Como la *maduración* de los sujetos será similar en ambos grupos (el *experimental* y el de *control),* las diferencias que en ellos se observen podrán atribuirse al efecto del *tratamiento,* y no a la evolución de los sujetos experimentales.

• *Historia*

Si la *maduración* hace referencia a la dinámica interna de desarrollo de las personas, la *historia* alude a acontecimientos externos que pueden alterar el curso del *experimento,* afectando a sus resultados.

Por *ejemplo,* durante el análisis del efecto de una campaña electoral en el voto de los electores, pueden acontecer hechos ajenos a la campaña que repercutan en la modificación del voto (como el conocimiento de un escándalo político protagonizado por uno de los líderes de los partidos en campaña electoral).

Al igual que la fuente de invalidez anterior, la *historia* es más factible en experimentos de larga duración. Asimismo, el recurso a un *grupo de control* ayuda al "control" de su influencia. Como su repercusión adquirirá igual relevancia en los distintos grupos de tratamiento, las diferencias observadas podrán atribuirse al *tratamiento* aplicado.

• *Instrumentación*

Cuando la medición de la variable *dependiente* se realiza varias veces (antes y después del *tratamiento),* las variaciones en las mediciones también pueden deberse a cambios producidos en el instrumento de medición (y no al efecto del *tratamiento).*

Por *ejemplo,* si se administra un *cuestionario* para comprobar los conocimientos previos de los alumnos antes de aplicarles el *tratamiento,* el *cuestionario* ha de ser similar en la primera y en posteriores mediciones. No deben existir variaciones importantes, ni en su contenido, ni en la forma de su administración.

Nuevamente, el recurso al *grupo de control* permite contrarrestar los efectos de alteraciones en los *instrumentos de medición.* Al ser el *instrumento de medición* igual en los distintos grupos, las diferencias en sus resultados se atribuirán al *tratamiento* dado.

• *Mortalidad experimental*

Durante el curso del *experimento,* algunos de los sujetos de la *muestra* de estudio pueden dejar de participar en la investigación. Su abandono (o *mortalidad experimental)* puede repercutir en los resultados de la investigación. Sobre todo, cuanta mayor disparidad exista en las características de aquellos que dejan el *experimento* y los que aún permanecen en él. Por lo que, la *muestra* final tendrá una composición distinta a la inicial, con la consiguiente pérdida de *representatividad.*

Obviamente, esta fuente de *invalidez* adquiere mayor alcance, conforme aumenta el período de duración de la investigación. Lo que favorece la aparición de *factores internos* (cansancio, pérdida de interés) y *externos* (cambio de residencia, de ocupación, o cualquier cambio en la forma de vida del sujeto que dificulte su continuidad en el *experimento),* que provocan la *mortalidad experimental.*

El uso de *varios grupos de tratamiento,* y la *comparación* de sus resultados, puede también ser de utilidad en el *control* de esta cuarta fuente de *invalidez.*

• *El efecto de la regresión*

Esta quinta amenaza a la *validez interna* se presenta cuando alguno de los grupos de *tratamiento* está integrado por sujetos que partieron de posiciones muy extremas en la variable *dependiente.* En una segunda medición, esta posición puede verse moderada, y no como consecuencia del *tratamiento,* sino de la tendencia observada en las personas de puntuar en posiciones menos extremas en mediciones posteriores.

Si se asignan los sujetos con posiciones iniciales extremas al azar a distintos grupos, las oscilaciones en sus puntuaciones (con independencia del *tratamiento)* será la misma en todos los grupos. Lo que posibilita la atribución de la disimilaridad de los resultados al efecto del *tratamiento.*

• *El efecto de la medición inicial*

El recurso al *pretest* (o medición de la variable *dependiente* previo al *tratamiento)* puede orientar a los sujetos experimentales sobre la finalidad del estudio, y de lo que

se espera de ellos. Éstos, como consecuencia, pueden modificar su actuación, afectando a los resultados de la indagación.

Por otra parte, el *pretest* también favorece la familiarización con el *instrumento de medición*. Esto contribuye a que el individuo mejore su puntuación, en una segunda medición, con independencia del *tratamiento* dado.

No obstante, la creación de un *grupo de control,* sometido también a mediciones previas y posteriores al *tratamiento,* contribuye a descartar esta fuente de *invalidez*. Como en todos los grupos ha habido una medición de la variable *dependiente* anterior al *tratamiento,* las diferencias que se observen tras la *manipulación experimental* podrán explicarse como resultados del *tratamiento*.

Estas seis *fuentes* importantes de *invalidez* se logran "controlar" con un *diseño experimental* que incluya la creación de un *grupo de control*. Pero la existencia de este grupo puede, a su vez, provocar otras *fuentes de invalidez*. Éstas se atribuyen a:

a) *Sesgos en la selección y asignación de los sujetos a los grupos de tratamiento* (experimental y de control). El perfil de los integrantes de cada grupo puede diferir antes de la aplicación del *tratamiento*. Lo que invalidaría los resultados atribuidos al efecto diferencial del *tratamiento*.

   Para obviar este tipo de *sesgo* se recomienda seguir procedimientos de selección aleatorios en la asignación de los individuos a los grupos (como ya se ha indicado).

b) *Compensación al grupo de control por la privación del tratamiento* (si éste se estima de valor). Esta *compensación* (un trato más favorable por parte del personal de la investigación, una mejor ubicación física del grupo u otra) resta eficacia al *grupo de control,* en la eliminación de explicaciones alternativas al *tratamiento*. En este caso, como en el anterior, el *grupo de control* que resulta no es "auténtico".

Asimismo, el conocimiento, por parte de los integrantes del *grupo de control,* de la privación de *tratamiento* que se les ha hecho, puede afectar su intervención posterior en la investigación:

a) *Desmoralizándose,* suspendiendo su actuación en el *experimento*.

b) *Rivalizando con el grupo experimental,* intentando su superación. Por *ejemplo,* si los alumnos del *experimento* anteriormente supuesto supiesen que se les ha privado, a propósito, de una práctica docente considerada "ideal", pueden desanimarse y/o irritarse bastante. El desánimo puede provocar su desatención en clase, con la consiguiente disminución en el rendimiento académico. Por el contrario, la irritación puede empujarles a la rivalidad y al deseo de competir con el *grupo experimental,* dedicando más horas a la preparación de la asignatura.

c) *Imitando el tratamiento dado al grupo experimental.* Los miembros del *grupo de control* pueden manifestar las mismas actitudes y conductas de los individuos tratados con el *estímulo experimental*.

En la medida en que el investigador prevea estas posibles *fuentes de invalidez interna,* podrá reducir la presencia de *explicaciones alternativas* a la *manipulación experimental*. Su investigación ganará en *validez interna*.

8.4. Los experimentos de campo frente a los de laboratorio

Los *experimentos* no sólo varían respecto al diseño, sino también dependiendo del contexto de la investigación. En general, se diferencian dos amplias modalidades de *experimentos:* los *experimentos de campo* y los de *laboratorio*.

Los *experimentos de campo* son aquellos que se llevan a cabo en contextos reales. o *naturales* (en contraposición a la *artificialidad* característica de los *experimentos de laboratorio).* Por ejemplo, en fábricas, en aulas escolares, en un parque público o donde quiera analizarse la reacción habitual o normal de la gente ante determinados hechos o conductas. El investigador puede, azarosamente, encontrarse con una determinada situación (por ejemplo, un accidente de tráfico), y observar cuál es la reacción de la gente (si socorren a las víctimas, huyen o únicamente observan impávidos). Pero, también él mismo puede crear la situación. Puede, por *ejemplo,* disponer una persona mal herida tumbada en una calle comercial para, a continuación, observar (como un transeúnte más) cuál es la reacción de la gente que pasa junto a la persona herida.

En los *experimentos de campo* no se realiza una selección aleatoria de las *unidades muestrales*. Tan sólo se observan a personas corrientes en su cotidianidad (en su hacer diario). Estas personas, por tanto, desconocen su participación en la investigación.

Asimismo, el propio entorno "natural", en que se desarrolla la investigación, dificulta el ejercicio del *control experimental a priori*. Lo que redunda en limitaciones a la *validez interna* del *experimento* (la posibilidad de eliminar *explicaciones alternativas* a las relaciones causales obtenidas), y mejoras en *validez externa* (la *representatividad* de la población observada, respecto al conjunto de la población). De ahí. la semejanza de los experimentos de *campo* con los *cuasi-experimentos* más que con *experimentos* propiamente dichos (Cook, 1983).

En cambio, los *experimentos de laboratorio* se presentan como los avales de los principios básicos de la *experimentación*. El investigador puede "manipular" cualquier situación que estime relevante a las hipótesis de su investigación. Él mismo controla las características que han de reunir las *unidades muestrales,* al igual que las singularidades ambientales que pudieran influir en los resultados del *experimento*.

A diferencia de los *experimentos de campo,* en los de *laboratorio* los sujetos suelen conocer su participación en el *experimento*. Ello puede provocar un *inconveniente* importante: la modificación de su conducta, el ajuste de ésta a lo que "se espera de él", y no a su proceder normal.

En 1962, Orne (en su famoso artículo "On the social psychology of the psychological experiment with particular reference to demand characteristics and the implications". en *American Psychologist,* 17) demuestra el *efecto de las características de la demanda* en los resultados de la investigación. Este autor argumenta cómo la identificación de los su-

jetos con el éxito del *experimento,* les lleva a ajustar sus conductas a lo que se espera de ellos, en su deseo de realizar una contribución "útil" a la investigación. Dicha contribución se entiende como corroboración de las *hipótesis* del estudio.

Pero no solamente las *características de la demanda* pueden afectar a la *validez interna* de los *experimentos de laboratorio*. También pueden incidir *características del experimentador*. En 1966, Rosenthal (en su conocida investigación *Experiment Effects in Research)* muestra cómo el experimentador puede convertirse en una fuente importante de *sesgo* en los *experimentos de laboratorio*. Características personales suyas (como el sexo, la edad, la étnia), junto a las singularidades de su actuación, pueden influir en los resultados de la investigación. En concreto, el investigador puede proporcionar a los sujetos experimentales señales que les indiquen si su reacción al *estímulo experimental* se adecúa a las expectativas de la investigación. Entre estas *señales* destacan el tono de voz y gestos (como asentir con la cabeza, girarla, cerrar los ojos, fruncir el ceño).

A estas *fuentes de invalidez interna,* se añaden otras actuaciones del experimentador, relacionadas con su función de observador y registrador de las respuestas a los *estímulos experimentales*.

En los *experimentos de laboratorio* también es común una fuente importante de *invalidez externa:* la mayoría de ellos se llevan a cabo con estudiantes universitarios, preferentemente de los primeros cursos de psicología. Orenstein y Phillips (1978) cifran entre un 73% y un 86% de los experimentos publicados en revistas científicas (dependiendo de la revista), el porcentaje de aquellos que utilizan a estudiantes como sujetos experimentales. Estos estudiantes a veces se ofrecen voluntarios, otras veces se les insta a ello (por cursar estudios de sociología o de psicología), y pueden o no percibir una remuneración por su participación en el *experimento*. Ello genera problemas en la generalización de los resultados *(validez externa),* debido a que estos estudiantes no representan al conjunto de la *población*. A veces, incluso, ni a la población estudiantil general.

La mayoría de los voluntarios son varones, con una necesidad elevada de aprobación social, y suelen ser más sociables y menos convencionales que aquellos estudiantes que no se ofrecen voluntarios (Orenstein y Phillips, 1978). De ahí la exigencia de comprobar las hipótesis experimentales en una amplia diversidad de sujetos (y no sólo con personal voluntario).

Además, la propia *artificialidad* creada en los *experimentos de laboratorio* (necesaria para el control de explicaciones alternativas) resta *validez externa* al *experimento;* es decir, la posibilidad de generalización de los hallazgos más allá de los confines del laboratorio. Esto provoca argumentos en contra de su utilidad, precisamente por la "artificialidad" del entorno donde éstos acontecen (Bryman, 1995).

8.5. Ventajas e inconvenientes de la experimentación como estrategia de investigación

Si por algo destaca la *experimentación,* es por su adecuación a proyectos de investigación que implica la comprobación de *hipótesis causales*. También se adapta al análisis de los procesos de interacción en grupos pequeños. Su coste suele ser inferior al de otras

estrategias de investigación (precisa menos personal y suelen analizarse pocos casos). A esta fundamental ventaja económica se añade la mayor facilidad de repetir un *experimento* en sujetos diferentes, en comparación con la *encuesta* (Babbie, 1992).

En su contra sobresale la referida "artificialidad" que supone la *experimentación* (específicamente, los *experimentos de laboratorio). Éste* es un grave *inconveniente* que limita las posibilidades de generalización de los resultados de la indagación.

A esta crítica se suman otras dos, relativas a la población y a las variables que se analizan. Pocas son las personas que desean ser conejillos de indias en una investigación. "Muchos adultos simplemente no están disponibles para participar en experimentos, y los voluntarios no son representativos de la generalidad de la población. Esto limita la investigación experimental a *poblaciones cautivas* (tales como estudiantes, delincuentes recluidos en instituciones correccionales, trabajadores de fábricas, miembros de las fuerzas armadas)" (Hakim, 1994: 102). En su mayoría, grupos pequeños y atípicos. Todo ello influye en la *representatividad* y posibilidad de generalización de los resultados de la indagación.

La *investigación experimental* también se halla circunscrita a temas específicos y a variables que puedan manipularse. Cuestiones éticas impiden la *manipulación experimental* de muchas variables de interés para el investigador social.

Por *ejemplo,* no sería ético separar a unos niños pequeños de su entorno para, con posterioridad, comparar su desarrollo socio-afectivo con el de aquellos niños que permanecieron en su propio entorno sociofamiliar. Tampoco sería admisible, desde el punto de vista ético, la asignación aleatoria de niños huérfanos a hogares caracterizados por experiencias de violencia en distintos niveles (con la finalidad de medir la influencia de los malos tratos en la posterior evolución conductual de los niños).

En el Cuadro 8.2 se resumen algunas de las ventajas e inconvenientes principales de la *experimentación* como *estrategia de investigación*.

CUADRO 8.2. Ventajas e inconvenientes de la experimentación como estrategia de investigación.

Ventajas	*Inconvenientes*
Permite la verificación de relaciones causales. Favorece el control *a priori* de las explicaciones alternativas a las relaciones observadas *(validez interna).* Ayuda al análisis de los procesos de interacción grupal. Facilidad de repetición de la investigación. Reducido coste económico de la investigación (en comparación con otras estrategias).	Problemas de inferencia y de generalización de los hallazgos de la investigación *(validez externa).* Las exigencias del control experimental limitan su aplicación a entornos artificiales (laboratorio), a grupos pequeños, y a unas cuantas variables. La utilización de personal voluntario (poblaciones cautivas) contribuye a la no representatividad de la muestra. Su materialización resulta difícil en entornos naturales. Efectos reactivos debidos a la presencia del investigador.

8.6. Ejemplo desarrollado de un experimento tomado de una investigación real

La investigación que a continuación se expone fue realizada en 1988 por E. Gaviria, J. Solano, M. Carmona y J. M. Fernández Dols: "Dominancia minoritaria: dos experimentos exploratorios", en *Revista de Psicología Social,* 1988, 3, 7-22.

El *objetivo* principal del estudio es el análisis de las relaciones de dominación (la relación entre mayoría y minoría). La hipótesis a comprobar es que, con independencia de los procesos de influencia social, el pertenecer a la minoría numérica implica ciertas ventajas posicionales, que pueden facilitar la dominación sobre el resto del grupo. Para ello se imponen dos condiciones previas:

a) Todos los sujetos deben ignorar (a lo largo de la sesión experimental) que son víctimas de un *tratamiento,* que les asigna un rol mayoritario o uno minoritario.
b) Los sujetos no deben haber interactuado previamente. La primera vez que interactúan es en la sesión experimental.

En esta investigación, la *sesión experimental* se divide en tres partes:

a) Unos minutos de interacción previos a la manipulación experimental.
b) La manipulación experimental propiamente dicha (que determina una mayoría y una minoría).
c) Tras dicha situación, un período de interacción posterior al tratamiento.

Por lo que, este *diseño experimental* se ajusta, en la notación de Campbell y Stanley (1970), al esquema de *diseño antes y después (O X O)*. La *hipótesis* se comprueba mediante dos *series experimentales:*

• *Experimento 1*

En primer lugar, se trata de comprobar si una situación de conflicto-divergencia entre mayoría y minoría implica una pauta de comportamiento competitivo, en la que emerge un sujeto dominante. En segundo lugar, se quiere determinar hasta qué punto el mero hecho de ser una minoría numérica facilita la dominación del individuo en el grupo.

A tal fin, se cita a 32 sujetos, 9 de los cuales abandonaron el *experimento* durante su realización. Por lo que fueron, al final, 23 los individuos observados. Se trataba de mujeres, estudiantes de los primeros cursos de universidad, seleccionadas aleatoriamente de las listas de secretaría. La *población* se distribuye, al azar, en 6 grupos de 4 miembros (excepto un grupo que fue de tres sujetos). Explícitamente se cuida que los

sujetos de la *muestra* no se conozcan entre sí, con el propósito de evitar que hubieran tenido una interacción previa.

Se crea una minoría de un sujeto (siempre el sujeto 4 de los grupos formados) frente a una mayoría de tres. La única característica del *tratamiento* que no era común para la mayoría y la minoría fue la designación de la minoría percibida como competente: en 4 casos, el experimentador asignaba la solución a la minoría numérica; en 2 casos, a la mayoría. En consecuencia, el *diseño* se caracteriza como un *diseño de grupo de control con tratamiento inverso y pretest/postest.*

Con respecto a la tarea del grupo, se idea un procedimiento que permite establecer un indicador de dominancia individual. Dicha *tarea postexperimental* consiste en construir una figura con siete piezas de un puzzle tipo Tangram (este puzzle permite un número prácticamente infinito de figuras).

Se seleccionan cuatro figuras (un barco, un hombre, un pájaro y un gato) consideradas de dificultad similar por parte de un equipo de jueces. Los 23 sujetos son adiestrados, unos días antes del *experimento,* en la realización de una de las figuras de modo que, una vez en grupo, cada sujeto sea capaz de realizar una figura distinta a la de los demás. Los sujetos no conocían esta circunstancia. Por el contrario, creían que todos habían sido entrenados para la misma figura. Tampoco conocían la relación entre el adiestramiento y el *experimento.* Momentos antes de la sesión experimental se comprueba que los sujetos recuerdan correctamente la construcción de la figura que se les había asignado. La finalidad era observar si un sujeto impone su solución unilateralmente a los otros integrantes del grupo.

En cambio, la *tarea pre-experimental* consiste en otro puzzle de distinto tipo; en este caso, los sujetos deben construir un polígono con un gran número de piezas, pero todos podían participar en la actividad simultánea y cooperativamente. Esta tarea permite verificar si un miembro del grupo impone una determinada solución, debido a que, durante la sesión experimental, los sujetos deben resolver el puzzle una sola vez y con un único conjunto de siete piezas.

Tras la sesión se lleva a cabo una *entrevista post-experimental,* en la que se solicita a los sujetos que indiquen: si conocían anteriormente a algún miembro del grupo; que intentasen recordar los juicios expresados por cada miembro del grupo durante la sesión experimental; que indicasen con cuál de los miembros del grupo estarían dispuestos a hacer una tarea similar; y, por último, que comentasen la tarea del grupo y las posibles causas de la divergencia.

Las dificultades para reclutar sujetos, enseñarles a resolver el Tangram, y mantenerlos aislados entre sí, hizo imposible la utilización de un tercer *grupo de control sin tratamiento.*

A lo largo de toda la sesión una cámara registra la interacción del grupo. En ningún caso se observa que los sujetos se pusieran de acuerdo para llevar a cabo una figura colectiva que no supusiera ninguna de las soluciones individuales. Siempre se observa una solución unilateral, y los sujetos minoritarios imponen su solución. Por tanto, no se produce ningún trabajo cooperativo. Siempre aparece un sujeto dominante, que coincide con el sujeto minoritario en tres de las seis sesiones. Mediante distintos *tests binomiales* se calcula la probabilidad de que sujetos mayoritarios y minoritarios impongan su solución al grupo.

Los hallazgos del *primer experimento* hacen necesario un *segundo experimento* que, desde un *diseño* similar, modifique las condiciones en el *tratamiento experimental* y las tareas *pre y post tratamiento,* con vistas a verificar estas observaciones.

• *Experimento 2*

Los *sujetos experimentales* son, en este caso, 24 estudiantes, de sexo femenino, que no habían interactuado previamente entre sí, seleccionadas al azar entre las personas que acuden al comedor universitario.

Para este *experimento* se siguió el mismo procedimiento general del *experimento 1*. Sin embargo, hay variaciones en el tipo de tareas *pre y post-tratamiento,* y en las variables manipuladas en dicho *tratamiento*.

El propósito del *experimento* es analizar la pauta de dominación minoritaria en un contexto cooperativo y ambiguo para observar, si en tales circunstancias, la minoría sigue mostrando una mayor tendencia a la dominación que la mayoría.

Para alcanzar dicho propósito, se repite el *diseño experimental 1*. Lo que permite comparar los resultados de ambos experimentos. Pero, en este caso, se intentan lograr unas relaciones de influencia mayoritaria-minoritaria más marcada, con la intención de comprobar el grado de interdependencia de dichos fenómenos con respecto al de dominación minoritaria.

Así, en un caso, la mayoría numérica es la que recibe el apoyo del experimentador en los ensayos divergentes, pero encontrándose, además, en una situación perceptiva carente de ambigüedad. En cambio, la minoría no sólo no recibe atribución de competencia, sino que también se halla en una situación perceptiva ambigua. En las sesiones experimentales del grupo con *tratamiento inverso* es la minoría la que se encuentra en situación perceptiva carente de ambigüedad frente a una mayoría en condiciones opuestas.

En este *experimento 2,* los sujetos trabajaron en grupo, antes y después del *tratamiento,* en un juego cooperativo. En concreto, construir una carretera de juguete o un pueblo en una mesa, utilizando materiales del *Test del Pueblo*.

La dominación de los sujetos aquí se establece de un modo menos artificioso que en el *experimento 1:* el experimentador indicaba la tarea y, una vez acabada (sentado todavía de forma aparentemente casual de cara a la pared, evitando toda interacción visual o facial con los sujetos), les solicitaba que explicasen lo que habían hecho.

A partir de los datos registrados en vídeo, se observa quién había actuado como portavoz de los miembros del grupo. Para ello se lleva a cabo una observación sistemática de la secuencia a intervalos de cinco segundos (registrándose diversas conductas verbales y no verbales de los sujetos).

De todos los registros producidos, había uno que indicaba quién dirigía la exposición por parte de los miembros del grupo. Se considera la frecuencia de *emisión de tanteos* como criterio más objetivo para detectar el habla acompañada por la dirección de la atención del grupo.

La *conducta de punteo* se registra en vídeo, analizándose la secuencia que comienza cuando el experimentador solicita al grupo que explique lo que ha hecho, y finaliza cuando los sujetos concluyen la explicación. Dicha secuencia se descompone en intervalos de cinco segundos. Dos observadores entrenados registran la presencia o ausencia de *punteo* en la conducta verbal de los sujetos del grupo.

Se aplica ANOVA *(análisis simple de la varianza),* con una transformación previa de los datos, obteniéndose un efecto significativo, y muy fuerte, en las diferencias entre mayoría y minoría para las observaciones *post-tratamiento*. Los ANOVA para las observaciones *pre-tratamiento* no son significativos en ningún caso. Lo que permite descartar la existencia de diferencias entre los grupos previas al *tratamiento*.

En suma, los resultados de los dos *experimentos* llevados a cabo demuestran la viabilidad de la *hipótesis* de la investigación. Parece que la aparición de la dominación no guarda relación directa con la influencia social, y que se da más fácilmente en los sujetos minoritarios.

Lecturas complementarias

Alvira, F. *et al.* (1981): *Los dos métodos de las ciencias sociales*. Madrid, CIS.
Arnau, J. (1984): *Diseños experimentales en psicología y educación*. México, Trillas.
Campbell, D. y Stanley J. (1970): *Diseños experimentales y cuasiexperimentales en la investigación social*. Buenos Aires, Amorrortu.
Cook, T. y Campbell, D. (1979): *Quasi-experimentation: design and analysis issues for fields settings*. Chicago, Rand McNally.
Morales Domínguez, J. *et al.* (1981): *Metodología y teoría de la psicología*. Madrid, UNED.
Saxe, L. y Fine M. (1981): *Social experiments. Methods for design and evaluation*. Beverly Hills, Sage.

Ejercicios Propuestos

1. ¿Qué tipo de experimento es el descrito en el apartado 8.6? ¿Cumple todos los criterios de validez? Añada otros diseños alternativos para cubrir el mismo objetivo de investigación.
2. Ilustre fuentes posibles de invalidez con un ejemplo.
3. Describa el desarrollo de un experimento publicado en una revista científica, desde sus inicios hasta su conclusión. Especifique si cumple los criterios de validez.
4. ¿Qué experimento diseñaría para comprobar la incidencia del tiempo atmosférico en el rendimiento académico? Señale posibles "variables perturbadoras" y cómo las controlaría.

CUARTA PARTE

EL ANÁLISIS Y LA PRESENTACIÓN DE LA INFORMACIÓN

9

EL ANÁLISIS DE LOS DATOS

Una vez que ha concluido la recogida de información, comienza la fase de *análisis e interpretación de los datos*. Ésta es una fase compleja e igualmente crucial en el proceso de investigación. Si bien, en los últimos años se ha visto beneficiada por los grandes avances producidos en el campo de la informática. En concreto, la expansión de paquetes estadísticos informáticos (adaptados a ordenadores personales), que cubren el amplio espectro analítico existente (tanto *cuantitativo* como *cualitativo)*. La implantación de estos programas ha adquirido tal relieve que, en la actualidad, sería inconcebible un *análisis de datos* ajeno al procesamiento informático.

En las páginas siguientes se ofrece una panorámica de las *técnicas cuantitativas de análisis de datos*. La notable pluralidad de éstas, unido a condicionantes de espacio, limitan su exposición a una mera descripción de rasgos distintivos de cada una de ellas. Para un mayor conocimiento y profundización, se remite a la consulta de bibliografía especializada en *análisis de datos* (véase la selección de textos que figuran al final del capítulo, y en la bibliografía general al término del libro).

9.1. Los preliminares del análisis de datos

"El paso más importante antes de comenzar a diseñar un proceso de entrada de datos es determinar qué programas se utilizarán para analizar los datos y convenciones concretas respecto a los formatos de ficheros y de *missing data,* que pueden manejarse para esos programas" (Fowler, 1988: 128). De ahí que la exposición de las *técnicas estadísticas de análisis* comience por sus *preliminares* esenciales: la creación de los *ficheros de datos,* junto a la depuración de la información (con especial referencia a la "no-respuesta").

9.1.1. La creación de los ficheros de datos

Cada programa informático comprende instrucciones concretas para la creación de *ficheros de datos* propios y *ficheros de definición de los datos*. No obstante, la mayoría de estos programas permiten el acceso a ficheros elaborados por otros programas.

Primero, se confecciona un *fichero de datos* para la introducción de la información obtenida durante la investigación. Los *ficheros de datos* presentan formas diversas, dependiendo del programa informático que se maneje. La más común es el fichero de datos en formato *ASCII (American Standard Code for Information Interchange)*. Este se compone de filas y columnas de números. Cada *fila* representa los datos recabados de cada sujeto o caso observado. En las *columnas,* en cambio, figuran los valores de las variables correspondientes a cada caso (véase el Cuadro 9.1). Este tipo de formato se precisa para ejecutar la mayoría de los *paquetes estadísticos genéricos* (como el SPSS, el BMDP o el SAS).

CUADRO 9.1. Extracto de un fichero de datos.

```
001211611450021402032540210011011934212416421102111005312123211200502013
002312911340010292016332221012101321122216221210211120143223612520202041
003811612430024332016341110611114300311111120000221130271124413211311122
004911291150021402032530200143221400222122111003132122232040207110002110
005211911450021212011141110923322113002221121121101001104422342614001103
006311811460731242014542210923411230012111211202002202432232211300310012
007611711440021212017541111042311230012211222200001002432236211040053212
```

Los *paquetes estadísticos genéricos* también pueden, sin embargo, leer *ficheros de datos* creados por *bases de datos* (como el DBASE) u *hojas de cálculo* (como LOTUS). Estos ficheros incluyen nombres para las variables, además de información relativa a la amplitud de la variable y la *codificación* de los valores de "no-respuesta" *(missing data).*

Cuando los datos se han registrado en *formato ASCII,* hay que elaborar también un *fichero de definición de datos*. Este comprende instrucciones precisas para la lectura de los datos que componen el *fichero de datos*. Estas instrucciones incluyen información primordial sobre las variables analizadas en la investigación: *columnas* donde se hallan ubicadas en el *fichero de datos;* las *etiquetas* dadas a las variables y a sus atributos; así como la especificación de cómo se *codifican* los valores *"missing"* (o sin respuesta), para su tratamiento en el análisis de los datos.

En el Cuadro 9.2 figura un extracto de un *fichero de definición de datos* en SPSS. Comienza con el nombre del fichero donde se han grabado los datos. Prosigue con la relación de variables y el número de las columnas donde éstas se encuentran registradas. Después, se asignan *etiquetas* a las variables que lo precisen, al igual que a sus valores. En el SPSS, como en los demás programas informáticos, se dan al usuario instrucciones

para la correcta elaboración de estos ficheros. También se expone cómo transportar y convertir ficheros de otros programas, su lectura y demás especificaciones necesarias para el análisis de los datos.

CUADRO 9.2. Extracto de un fichero de definición de datos en SPSS.

```
Data list file= 'menores.dat'.
/id 1-7 anio 8 sexo 9 edad 10 nacional 11 etnia 12 anorma 13 alcohol 14
heroina 15 cannabis 16 estudio1 to estudio2 17-20 desescol 21 profepa 22
situacpa 23 apreviv 24 ambiente 25 fuga 26 compania 27.
Variable labels desescol 'desescolaridad'
/profepa 'profesion del padre'
/situacpa 'situación empleo padre'
/apreviv 'apreciación vivienda'.
Value labels sexo 1 'varon' 2 'mujer'
/nacional 1 'espanol' 2 'extranjero'
/estudio1 to estudio2 01 'primero EGB' 02 'segundo EGB' 03 'tercero EGB'
04 'cuarto EGB' 05 'quinto EGB' 06 'sexto EGB' 07 'septimo EGB'
08 'octavo EGB' 09 'FP' 10 'BUP' 11 'compensatoria' 12 'otros'
13 'analfabeto'.
Missing values anio to compania (0).
Frecuencies anio to compania /statistics=all /hbar.
Save outfile='menores.sys'.
```

Aparte de estos ficheros, pueden crearse *subficheros* específicos con objeto de facilitar los análisis. Especialmente, cuando el tamaño de la muestra es elevado y el *fichero de datos* originario adquiere un gran volumen.

Estos *subficheros* incluyen una *muestra de variables,* restringida a aquellas que sean de interés para la ejecución de análisis concretos. También, pueden representar una submuestra aleatoria de los casos observados en la investigación. En este último caso, la finalidad no sería tanto el ahorro de tiempo en el análisis de la información, sino la *validación* de los resultados estadísticos obtenidos de la otra submuestra de la muestra global. Así, por *ejemplo,* para la *validación* de *análisis multivariables* (como el de *regresión),* se recomienda la división de la *muestra* en dos *submuestras:* una, para la consecución del modelo; y la otra, para su *validación.*

9.1.2. La depuración de la información

A la creación de ficheros sigue la *depuración de los datos,* como antesala del análisis. El investigador ha de identificar posibles *errores* cometidos en la *grabación de los datos.* A tal fin resulta conveniente solicitar al programa la relación de *frecuencias* de todas las *variables* introducidas en el *fichero de datos ("frequencies"* en SPSS, o *"univariate"*

en SAS, por ejemplo). Esta relación incluye un listado con todos los *valores* de cada *variable*, la *frecuencia (absoluta y relativa)* de cada uno de ellos, los casos sin respuesta *(missing data), estadísticos univariados* y las *representaciones gráficas* que se soliciten.

De esta relación, se observará si existen anomalías en los *valores* de las *variables* codificadas. Más concretamente, si alguna de ellas incluye valores ajenos al recorrido o *rango* definido de la *variable*.

Por *ejemplo*, si en la variable sexo, que se ha codificado con sólo dos opciones de respuesta (1 'varon' 2 'mujer'), aparecen casos con valores superiores a dos (3, 4, 5 u otro), éstos corresponderían a sujetos que han sido erróneamente codificados o grabados en el ordenador. Por lo que, habría que proceder a su comprobación y corrección posterior.

Algunos programas informáticos (como el SAS o el SPSS) proporcionan, además, especificaciones dirigidas a la *depuración de los datos* (tanto durante su introducción, como una vez concluida ésta).

Si se observan incongruencias en los *valores* de las *variables* (como la anteriormente ejemplificada), ha de procederse a su localización y corrección. Los *errores* pueden deberse a fallos en la introducción de los datos en el ordenador, pero no siempre. También pueden ser ocasionados por deficiencias en la recogida de información.

En el primer caso, la corrección resulta más viable e inmediata: se revisan los *cuestionarios* (u otro instrumento de recogida de datos que se haya empleado), hasta localizar los casos en que se han grabado mal los *valores* de las *variables;* posteriormente, se introducirían los *códigos* correctos de las *variables* correspondientes.

Pero, si los *errores* se deben a un mal registro de la información en el instrumento de medición, las posibilidades de corrección se restringen. La dificultad de contactar de nuevo con la fuente de información (las *unidades muestrales)* lleva, no a la transformación de *códigos numéricos,* sino a la eliminación de aquellos casos con datos incorrectos o inconsistentes. Éstos se sumarían a aquellos que originariamente no proporcionaron información alguna. De esta forma, se incrementaría el volumen de los llamados *"missing values"* (o *valores* con los que se codifican las respuestas en blanco o incorrectas, y que se dan por *perdidos).*

Cuando unos casos específicos presentan muchos *"missing values"* en la mayoría de las *variables,* suele decidirse su exclusión del *fichero de datos* (salvo que al investigador le interese el análisis y descripción de estos casos). Igualmente, si de alguna variable se obtiene escasa información (teniendo un elevado porcentaje de valores *missing),* también suele optarse por su exclusión para el resto de los análisis.

En general, antes de proceder al *análisis de los datos,* el investigador evalúa los *porcentajes de respuesta* (para cada *variable),* y los *"outliers"* registrados en la *matriz de datos.*

Por *"outliers"* se entiende cualquier observación o caso que muestre inconsistencia con la serie global de datos. Su identificación requiere la realización de *análisis univariables* (para cada una de las *variables),* tanto numéricos como gráficos. Además, la mayoría de los programas informáticos incluyen instrucciones específicas para la detección de *"outliers".*

En cuanto a la *"no-respuesta",* su evaluación resulta igualmente exigida. Diversos autores, como Bourque y Clark (1994), recomiendan la *comparación* de las caracte-

rísticas demográficas de la *muestra* con las correspondientes a la *población* de la que ésta procede. Para ello se emplea el *Censo de Población,* u otra fuente de datos estadísticos o estudio, que describa al conjunto de la *población.*

Si de este análisis se dedujese la no *representatividad* de la *muestra,* el investigador deberá establecer la magnitud de las diferencias entre la *población* y la *muestra.* Esto es importante para la delimitación de las posibilidades de *inferencia* de las *estimaciones muestrales.* En palabras de Arber (1993: 71):

> "La capacidad para realizar inferencias de una muestra a una población se basa en el supuesto de que la muestra lograda no esté sesgada por la *no-respuesta.* En la medida en que aquellos que no responden difieran de forma significativa de aquellos que sí responden, el investigador tiene una muestra sesgada."

9.2. El análisis estadístico univariable

En el diseño de la investigación ya se prevén los análisis a realizar con la información reunida en el desarrollo de la investigación. Aunque ha de matizarse que el *proyecto de análisis* no es inmutable. Depende, en gran parte, de la *cantidad* y *calidad* de los datos que se recaben. De ahí la importancia que adquieren, en cualquier indagación, los *análisis exploratorios,* como paso exigido y previo a la decisión de qué técnica analítica *(bivariable* y *multivariable)* se va a aplicar.

9.2.1. La distribución de frecuencias

En la *exploración de los datos,* primero se procede a un análisis exhaustivo de cada *variable* incluida en la *matriz de datos (análisis univariable).* Para cada una de las *variables* se calcula su *distribución o tabla de frecuencias.* La *tabla de frecuencias* –como puede verse en el Cuadro 9.3– incluye los distintos *valores* que presenta la *variable* (distribuidos en *clases* o *categorías),* acompañados por su *frecuencia* (es decir, el número de veces en que aparecen).

En la primera columna (encabezada por el rótulo *"value label")* figuran los distintos atributos que componen la *variable.* La siguiente columna *("value")* muestra el *valor* dado a cada *atributo.* En la tercera columna (la denominada *"frecuency")* se hallan las *frecuencias absolutas;* o sea, el número de casos (de la *muestra)* que comparten cada uno de los *valores* de la *variable.*

Para conocer la importancia de cada *valor,* y a efectos comparativos, se obtienen las *frecuencias relativas o porcentuales,* que representan cada *valor* en el conjunto de la *muestra* (columnas 4.ª y 5.ª). Primero, se calculan los *porcentajes* para toda la *muestra;* segundo, exclusivamente para aquellos casos que han proporcionado información al respecto (columna de *"valid percent"* o porcentaje válido). En esta columna no se consideran, por tanto, los *"missing values".*

Este desglose de casos, en función de si aportan o no información, permite conocer la proporción de *"no-respuesta"* de cada *variable.* Este conocimiento adquiere especial relevancia para posteriores análisis.

Por último, se calculan las *frecuencias relativas acumuladas ("cum percent"),* a partir de las *frecuencias* contenidas en la columna 5.ª. Estas *frecuencias acumuladas* denotan la proporción de casos (válidos) que se encuentran por debajo, o por encima, de un determinado *valor* de la *variable.*

Cuando la *variable* está medida a nivel de *intervalo,* se aconseja la previa *agrupación* de los *valores.* Ello facilitará su presentación en una *tabla de frecuencias* de menores dimensiones. A este respecto, algunos autores –como Bryman y Cramer (1995)– sugieren que el número de *categorías* diferenciadas esté comprendido entre 6 y 20. Argumentan que menos de 6 y más de 20 categorías pueden distorcionar la *forma* de la distribución de la variable.

Los *estadísticos univariables* que figuran en el Cuadro 9.3 se comentan en la sección 9.2.3, dedicada a su exposición.

CUADRO 9.3. Ejemplo de una tabla de frecuencias para la variable "estudio" en SPSS.

Value label	Value	Frecuency	Percent	Valid Percent	Cum Percent
Primero EGB	1	8	1.1	1.1	1.1
Segundo EGB	2	8	1.1	1.1	2.2
Tercero EGB	3	31	4.3	4.3	6.6
Cuarto EGB	4	59	8.3	8.3	14.8
Quinto EGB	5	119	16.6	16.7	31.5
Sexto EGB	6	152	21.3	21.3	52.8
Septimo EGB	7	177	24.8	24.8	77.6
Octavo EGB	8	72	10.1	10.1	87.7
FP	9	27	3.8	3.8	91.5
BUP	10	7	1.0	1.0	92.4
Compensatoria	11	1	.1	.1	92.6
Analfabeto	13	53	7.4	7.4	100.0
.	0	1	.1	MISSING	
	TOTAL	715	100.0	100.0	

Primero EGB 8
Segundo EGB 8
Tercero EGB 31
Cuarto EGB 59
Quinto EGB 119
Sexto EGB 152
Septimo EGB 177
Octavo EGB 72
FP 27
BUP 7
Compensatoria 1
Analfabeto 53

Mean	6.566	Std Err	.091	Median	6.000
Mode	7.000	Std Dev	2.420	Variance	5.856
Kurtosis	1.644	S E Kurt	.183	Skewness	1.000
S E Skew	.091	Range	12.000	Minimum	1.000
Maximum	13.000	Sum	4688.000		
Valid Cases	714	Missing Cases	1		

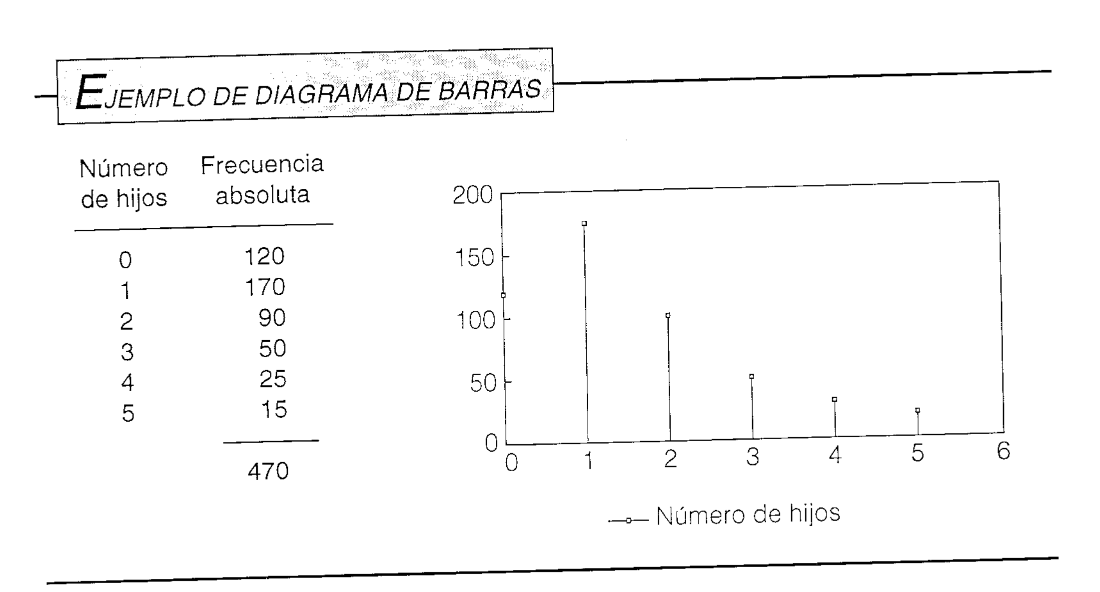
9.2.2. Representaciones gráficas

La información contenida en una *tabla de frecuencias* también puede representarse de forma gráfica. Ello ayuda a la visualización global de la concentración, o dispersión, de los datos en la variable considerada.

Dos de los *gráficos* habituales en la representación de *frecuencias* son el *diagrama de barras* y el *histograma*. A ellos se suman otros también usuales en el *análisis exploratorio,* como el diagrama de *"tronco y hoja"* y la *"caja"*; o el *polígono,* las *ojivas,* y el *gráfico de sectores,* entre la amplia variedad gráfica existente.

• *Diagrama de barras*

Consiste en una serie de "barras" (una para cada categoría de la variable), cuyas longitudes expresan las *frecuencias* de cada *atributo* de la *variable.*



Número de hijos	Frecuencia absoluta
0	120
1	170
2	90
3	50
4	25
5	15
	470

• *Histograma*

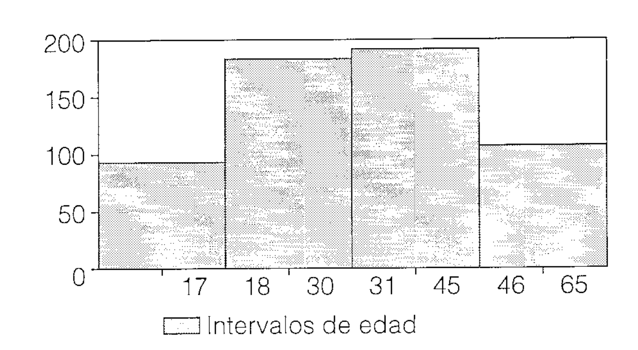
Un gráfico de contenido similar al anterior, aunque más adecuado a la representación de variables *métricas* (de *intervalo* o de *razón).*

En lugar de barras, el gráfico se compone de *rectángulos* unidos. Razón por la que se exige que la variable esté agrupada en *intervalos*. Estos forman las bases de cada uno de los rectángulos, que están delimitados por los *límites (reales)* de los respectivos *intervalos*. En cambio, la altura de los rectángulos es proporcional a la *frecuencia (absoluta* o *relativa)* de cada *intervalo*.

La suma total de las áreas de los rectángulos será igual a 1 (dado que la suma de todas las proporciones es la unidad).

EJEMPLO DE HISTOGRAMA

Edad	Frecuencia absoluta
Menos de 18	90
18 – 30	170
31 – 45	185
46 – 65	115
	560



• *El "Tronco y las Hojas" ("Stem-and-Leaf")*

Constituye un gráfico parecido al *histograma,* pero integrado por los *dígitos* de los *valores* de las *variables.* Los *dígitos* se dividen entre dos. Los situados a la izquierda del punto (el *tronco)* figuran ordenados verticalmente, en orden creciente (de arriba a abajo). Por el contrario, los *dígitos* a la derecha del punto (las *hojas),* se disponen horizontalmente, aunque también en sentido creciente (de menor a mayor).

El *dígito* a la izquierda (la *columna)* que comprenda más *valores* a la derecha será aquél en el que se agrupen un mayor número de casos en la distribución. Por esta y otras razones, este gráfico suele tomarse como referente de las medidas de *tendencia central* de una distribución de frecuencias (a las que se hará referencia en el apartado 9.2.3).

EJEMPLO DEL GRÁFICO "EL TRONCO Y LAS HOJAS"

En el siguiente gráfico, puede observarse que la *fila* tercera representa los valores de mayor frecuencia en la distribución. En concreto, los *valores* 36, 36, 37, 37, 37, 38, 38, 39, 39, 39.

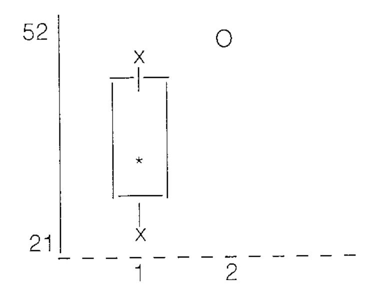
```
2 • 1 2 3 4
3 • 0 1 2 2 3 3 5
3 • 6 6 7 7 7 8 8 9 9 9
4 • 1 1 1 5 6 7 7
5 • 2 2
```

• *La "Caja" ("Box-plot o Box-and-Whisker")*

Ofrece una visión global de la distribución, más sintética que la anterior. En ella la *variable* se representa de forma que el 50% de los casos queden comprendidos en el interior de la "caja". En los extremos se sitúan, respectivamente, el 25% superior e inferior.

Con un asterisco se señala la *mediana;* con una "X" los valores máximos y mínimos; y con una "O", los *"outliers"*. De esta manera se proporciona (gráficamente) información referente a la *mediana,* el *primer cuartil* (el 25% de los casos iniciales), el *tercer cuartil* (el 25% finales), y el *recorrido intercuartílico* (el 50% de los casos centrales) de la distribución de frecuencias. Ello exige que el nivel de medición mínimo de la variable sea el *ordinal.* En caso contrario, no podría estimarse el valor de la *mediana*, ni de ningún estadístico que precise de la ordenación de los *valores* de la variable, en un sentido creciente o decreciente (los *cuantiles).*

EJEMPLO DEL GRÁFICO LA "CAJA"



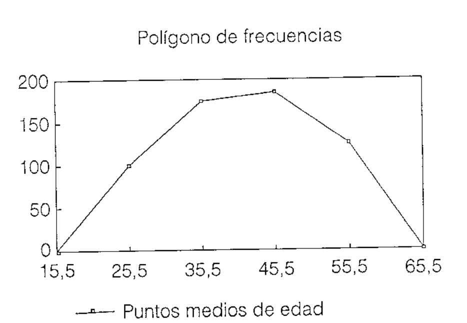
• *Polígono*

Un gráfico lineal que se traza sobre los *puntos medios* de cada *intervalo*, a una altura proporcional a su frecuencia *(absoluta* o *relativa).* Por lo que, de su visualización también se extraerán los *valores* con mayor (y menor) *frecuencia* en una distribución.

Los *puntos medios* se calculan sumando los *límites* de los *intervalos,* y dividiéndolo entre 2. De esta forma se obtiene el *valor* que representará al *intervalo* a todos los efectos. Por ejemplo, (21 + 30)/2 = 25,5.

EJEMPLO DE POLÍGONO DE FRECUENCIAS

Edad	Frecuencia absoluta	Punto medio
21 – 30	90	25,5
31 – 40	170	35,5
41 – 50	185	45,5
51 – 60	115	55,5
	560	



- *Ojivas*

Polígonos de frecuencias acumuladas que muestran la *frecuencia* de casos por encima, o por debajo, de un determinado *valor* de la distribución.

La *ojiva* será *"menor que"*, si se consideran los casos que hay por debajo de un *valor*. Por el contrario, será *"mayor que"*, cuando se representan los casos que comparten un *valor* superior de la distribución.

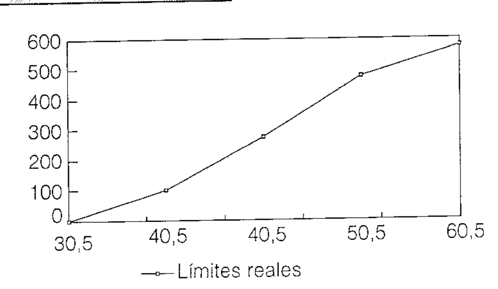
En el eje horizontal (de *abscisa)* se sitúan los *límites reales* de los *intervalos* de la *variable,* mientras que en el eje vertical (de *ordenada)* se disponen las *frecuencias acumuladas (absolutas* o *relativas)* de cada *intervalo.*

Para el cálculo de las *frecuencias acumuladas* se tiene en cuenta si interesa conocer el número (o la proporción) de casos que hay por debajo *(ojiva "menor que"),* o por encima *(ojiva "mayor que")* de un *valor* específico de la distribución.

EJEMPLO DE OJIVAS "MENOR QUE" Y "MAYOR QUE"

Ojiva "menor que"

Edad	Frecuencia absoluta	Frecuencia acumulada
21 – 30	90	90
31 – 40	170	260
41 – 50	185	445
51 – 60	115	560
	560	



Ojiva "mayor que"

Edad	Frecuencia absoluta	Frecuencia acumulada "más de"
21 – 30	90	560
31 – 40	170	470
41 – 50	185	300
51 – 60	115	115
	560	

600
500
400
300
200
100
0
30,5 40,5 40,5 50,5 60,5
Límites reales

• *Gráfico de sectores (o en forma de "tarta")*

A diferencia de los gráficos anteriores, éste se representa mediante un círculo, dividido en "sectores", cuyos *ángulos* indican el porcentaje de casos que comparten cada atributo de la variable.

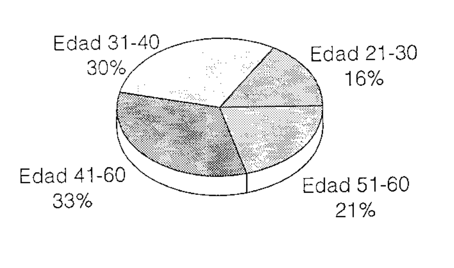
Para su obtención manual, se multiplican las *frecuencias relativas* (porcentual) de los *intervalos* (o valores) de la variable por 3,6 (que son los *grados* que corresponden a cada unidad porcentual: 360/100). De esta forma se transforman los *porcentajes* en *grados,* lo que posibilita su representación en una circunsferencia.

No obstante, el empleo de programas gráficos (como el Harvard Graphics) hace innecesaria la realización de estos cálculos. Basta con conocer la *frecuencia* de cada *valor* de la distribución.

Los "sectores" del gráfico suelen representarse con colores o trazas diferentes, que proporcionen un mayor contraste entre ellos.

EJEMPLO DE GRÁFICO DE SECTORES

Edad	Frecuencia absoluta	Frecuencia relativa (%)	Grados
21 – 30	90	16	58
31 – 40	170	30	108
41 – 50	185	33	119
51 – 60	115	21	75
	560		360



Para un conocimiento más amplio de éstas y demás representaciones gráficas comunes en la investigación social, se remite a textos específicos sobre *Gráficos* (como el de Alaminos (1993), y otros citados en la bibliografía).

9.2.3. Estadísticos univariables

Aparte de los *gráficos,* en la *exploración univariable* también se emplean estadísticos para medir, de manera precisa, la distribución de los valores de una variable (véase Cuadro 9.3). Su uso dependerá, sobre todo, del nivel de *medición* de la *variable.*

Los *estadísticos univariables* se agrupan en tres grandes apartados:

a) Medidas de tendencia central.
b) Medidas de dispersión.
c) Medidas de forma de la distribución.

• *Medidas de tendencia central*

Describen cómo se agrupan los *atributos* de una *variable* alrededor de un "valor típico" de la distribución. Por lo que, proporcionan una síntesis de la información contenida en la distribución.

Las medidas de *tendencia central* más empleadas en la investigación social son las siguientes: la *media,* la *mediana* y la *moda.*

1) La *media* es la medida más representativa, siempre y cuando la variable sea cuantitativa (de *intervalo* o de *razón).* Su cálculo precisa de la participación de todos los valores de la distribución. Cada uno de ellos se multiplica por sus respectivas *frecuencias absolutas*. Después, se suman todos los productos, y el resultado se divide por el número total de casos. De esta forma se obtiene el *promedio* de los valores de la distribución, que es como se define la *media aritmética.* El *inconveniente* fundamental de este estadístico es que se ve afectado por *valores* muy extremos en una distribución. Cuando la variable incluye *valores* muy dispares (o extremos), la *media* pierde representatividad a favor de la *mediana,* como *medida de tendencia central.*
2) La *mediana* es el valor que divide a la distribución en dos partes iguales. Se sitúa en el medio de la distribución. Lo que permite el conocimiento de los *valores* de mayor representación en la *muestra.* Su cálculo precisa que la variable sea, al menos, ordinal.
3) La *moda* denota el valor de mayor frecuencia en una distribución; aquel que más casos comparten. La distribución puede ser *unimodal* (una sola *moda),* pero también *bimodal* (dos *modas),* o *multimodal* (más de dos *modas).* Esto dificulta su interpretación, en caso de coexistir varias *modas.* A este *inconveniente* primordial se suma otro importante: en su cálculo no intervienen todos los *valores* de la distribución. Si bien, su principal *ventaja* es su universalidad. Pue-

de estimarse para cualquier tipo de variable, ya que el *nivel de medición* mínimo exigido es el *nominal.*

Además de estos estadísticos, existen los *cuantiles* como *medidas de tendencia no central.* Éstos representan *valores* que dividen a la distribución en partes iguales. Los *cuartiles* la dividen en cuatro partes iguales (cada una de ellas incluye al 25% de los *valores);* los *deciles* en diez partes; y los *percentiles* en cien partes. Su cálculo se asemeja al de la *mediana.* De hecho, el *cuartil* dos, el *decil* cinco, y el *percentil* cincuenta, expresan el *valor mediano.*

• *Medidas de dispersión*

Al conocimiento de los *valores centrales* de la distribución le sigue la medición de su representatividad: la mayor o menor *variabilidad* existente en torno a la *media* o la *mediana* de la distribución.

Las medidas de *dispersión absolutas* más comunes son el *rango* (o recorrido), la *desviación típica,* y la *varianza.*

1) El *rango o recorrido* expresa el número de valores incluidos en la distribución. Estos se obtienen de la diferencia entre el valor superior y el inferior. Su comprensión es sencilla, aunque presenta, en su contra, un *inconveniente* importante: es susceptible a la distorción proporcionada por valores muy extremos en una distribución. Ello se debe a la única consideración del primer y último valor de la distribución.

   Para solventar esta deficiencia, suelen aplicarse otras medidas de *rango,* que emplean un mayor volumen de información. Si bien, éstas exigen que la variable sea, al menos, *ordinal.* Se trata de los *recorridos: intercuartílico* (la diferencia entre el tercer *cuartil* y el primero; por lo que incluye al 50% de los valores centrales de la distribución), *semiintercuartílico* (el 25% de los valores centrales), *interpercentílico* (el 80% de los valores centrales, al ser la diferencia entre el *percentil* 90 y el 10), y el *semiinterpercentílico* (el 40%).
2) La *desviación típica* es el promedio de la desviación de los casos con respecto a la *media.* Como indicador de heterogeneidad (o de dispersión de los valores de una distribución), su estimación se exige siempre que se calcule la media, porque ayudará a la interpretación de su representatividad en la distribución.

   Al igual que la *media,* su valor viene expresado en la unidad de medición de la *variable,* y únicamente puede calcularse cuando la variable es *cuantitativa.*
3) La *varianza* constituye otra medida de heterogeneidad de una distribución. Se define como el cuadrado de la *desviación típica.* Su valor expresa el grado de heterogeneidad de una población respecto a la *variable* medida, siendo sus características similares a la *desviación.*

A partir de la *desviación*, puede calcularse una medida de *dispersión relativa* que ayudará en su interpretación. Se trata del *coeficiente de variabilidad de Pearson*. Una medida estandarizada que se obtiene del cociente entre el valor de la *desviación* y la *media aritmética*. Su valor se expresa en porcentajes, siendo de utilidad en la comparación de la homogeneidad de dos o más grupos respecto a una o más variables.

Por último, cuando se calcula la *mediana*, también suelen estirmarse estadísticos que midan su representatividad en la distribución. Concretamente, la *desviación media de la mediana* y el *coeficiente de variabilidad de la mediana de Pearson* (obtenido del cociente entre la *desviación media de la mediana* y la *mediana)*. Su interpretación es análoga a los estadísticos anteriores.

Para un conocimiento más detallado, remito a la consulta de cualquier manual de estadística.

• *Medidas de la forma de la distribución*

Este término comprende la disposición espacial de los *valores* en una distribución: su *asimetría* y *curtosis*.

1) La *asimetría* constituye un indicador de la *agrupación* de las *frecuencias* en la *curva* de una distribución. Cuando su valor es cero, expresa que la *curva* es *simétrica;* es decir, que coinciden los valores de la *media,* la *moda* y la *mediana*. Si es mayor de cero, significa que la distribución es *asimétrica a la derecha (o de sesgo positivo)*. Los casos se agrupan a la izquierda de la *curva*. Esto significa la mayor presencia en la distribución de *valores* inferiores a la *media*.

   En cambio, si el valor del coeficiente de sesgo es inferior a cero, denota que la agrupación de los *valores* se produce a la derecha de la *curva*. Por lo que habría una mayor representación de los *valores* superiores a la *media*.

   En general, un valor superior a 0.8 (positivo o negativo) indica que la *asimetría* de la variable es importante.

2) La *curtosis* hace referencia a la mayor o menor *concentración* de *valores* en torno a la *media* de la distribución. Si existe una elevada concentración, la distribución será *leptocúrtica*. Esto significa que la *media* es muy representativa, al haber muy poca dispersión respecto a ella.

   La distribución será *platicúrtica,* en el caso contrario: cuando existe una elevada *dispersión* de *valores* respecto a la *media*. Lo que expresa su escasa representatividad.

Con estos estadísticos lo que se pretende es comprobar si existen valores muy extremos en una distribución. En caso de existir, puede decidirse su transformación, con la finalidad de alcanzar una distribución que se aproxime a la *normal (simétrica* y *mesocúrtica)*.

Tras los primeros *análisis univariados,* el investigador adquiere una descripción inicial de cada una de las *variables* que componen la investigación. Este conocimiento le

puede llevar a introducir *modificaciones en la configuración de las variables,* de forma que se faciliten los análisis posteriores. Estas *modificaciones* consisten, esencialmente, en agrupaciones de *variables* o de sus *atributos*.

Cuando en una distribución se observa que una o varias *categorías* presentan una *frecuencia* muy baja, habría que proceder a su *agrupación* con otras *categorías* de similar significado (si se pretende, con ellas, efectuar *análisis bivariables* o *multivariables).* Si la variable es *nominal,* la agrupación se produce con *categorías* que presentan un nexo común. Pero, si la variable es *ordinal,* de *intervalo,* o de *razón,* han de agruparse sólo los *valores* que se hallen más próximos en la escala (por *ejemplo,* las categorías "extrema derecha" con "derecha").

A la nueva *categoría* (formada de la agrupación de dos o más categorías originales) habría que darle una nueva denominación, que resuma la variedad de *atributos* que comprende. De este modo las posibilidades de análisis, de variables con *atributos* de escasa representación en la *muestra,* se amplían; aunque la agrupación lleve consigo pérdida de información.

Igualmente, puede *agruparse variables* similares, con la finalidad de componer una medida única que sintetice la información contenida en *variables* análogas.

9.3. El análisis bivariable

Después de la realización de los *análisis exploratorios,* procede la realización de *análisis bivariables,* tanto con fines *descriptivos* (describir al conjunto de la población observada), como *explicativos* (analizar posibles relaciones causales entre dos variables: la independiente y la dependiente).

9.3.1. Las tablas de contingencia

En la investigación social, la práctica habitual es la confección de *tablas de contingencia*, formadas del cruce, al menos, de dos *variables*. Estas *tablas* generalmente se obtienen mediante los comandos CROSSTABS y TABLES, en la mayoría de los programas estadísticos.

En estos comandos se especifican las *variables* a cruzar. Si puede establecerse una *relación causal* entre ellas, el orden convencional de exponer las *variables* es: *dependiente* "by" *independiente*. De esta forma, la *variable dependiente* figurará en las *filas,* mientras que la *independiente* en las *columnas*.

Si se desean introducir *variables de control* (para la eliminación de *explicaciones alternativas),* estas terceras y, a veces, cuartas variables, se añaden a las anteriores, siendo igualmente precedidas por la preposición "by" (por *ejemplo,* CROSSTABS = EDAD BY HERMANOS; CROSSTABS = MEDIDA BY DELITO BY SEXO). No obstante, en el programa informático que se emplee, se especifican las instrucciones propias para la ejecución del comando.

Además de las *variables,* ha de delimitarse la información que se precise: *frecuencias absolutas, porcentajes* (horizontales, verticales, totales), y *estadísticos* que midan el grado y la significatividad de la relación entre las *variables.*

El Cuadro 9.4 muestra una *tabla de contingencia simple* obtenida mediante el comando CROSSTABS de SPSS. En él se solicitaron las *frecuencias absolutas ("count"),* y las *relativas:* porcentajes horizontales *("row")* y verticales *("column"),* para cada *casilla* de la *tabla.* A ello se sumaron los *estadísticos de contingencia.*

CUADRO 9.4. Una tabla de contingencia simple mediante SPSS.

Crosstabulation: Delito
By medida

Medida Delito	Count Row Pct Col Pc	Amonestación 1	Libertad vigilada 2	Internamiento 3	Row total
Hurto	1	40 51.3 19.3	26 33.3 11.7	12 15.4 12.2	78 14.8
Intimidación con armas	2	12 19.7 5.8	36 59.0 16.1	13 21.3 13.3	61 11.6
Robo sin intimidación	3	116 45.8 56.0	93 36.8 41.7	44 17.4 44.9	253 47.9
Insumisión paterna	4	10 13.7 4.8	43 58.9 19.3	20 27.4 20.4	73 13.8
Otras infracciones	5	29 46.0 14.0	25 39.7 11.2	9 14.3 9.2	63 11.9
	Column total	207 39.2	223 42.2	98 18.6	528

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
41.32788	8	.0000	11.322	None

Statistics	Symmetric	With DELITO Dependent	With MEDIDA Dependent
Lambda	.07069	.00000	.13443
Uncertainty Coefficient	.03488	.03033	.04102
Somers'D	.04951	.05219	.04710
Eta		.07227	.17561

Statistics	Value	Significance
Cramer's V	.19783	
Contingency Coefficient	.26943	
Kendall's Tau B	.04958	.0952
Kendall's Tau C	.04959	.0952
Pearson's R	.04051	.1764
Gamma	.07346	

Number of Missing Observation = 0

Cada *casilla* es el resultado del cruce de una *fila* con una *columna* (es decir, del cruce de un *atributo* de una *variable* con el *atributo* de la otra *variable).*

La lectura de las *tablas* con frecuencia se limita a comentarios porcentuales. Se contrastan los *porcentajes* de cada *casilla* para comprobar la existencia de variaciones entre los distintos *atributos* de las *variables*. A tal fin se calculan los *porcentajes* a partir de los *marginales* de la *tabla*.

Si se toma como base el total de *filas,* el *porcentaje* será *horizontal,* y las comparaciones porcentuales (entre los subgrupos) se efectuarán verticalmente. Por el contrario, cuando la base la constituye el total de *columnas*, se procede a la inversa: el *porcentaje* será *vertical,* y las comparaciones de porcentajes en sentido horizontal.

El investigador deberá escoger entre uno u otro tipo de *porcentaje* (horizontal o vertical), en conformidad con los *objetivos* del estudio y las *hipótesis* que compruebe. Si bien, existen mayores restricciones en *estudios explicativos* que en los *descriptivos*.

Si la finalidad de la investigación es la búsqueda de *relaciones causales,* los *porcentajes* se estiman sólo en el sentido de la variable *independiente*. Esta variable suele situarse en las *columnas,* y la *dependiente* en las *filas;* salvo que el elevado número de *atributos* de la variable *independiente* desaconseje su ubicación en las *columnas*. Esta disposición de las variables responde a la mayor facilidad de lectura (en la cultura occidental) en sentido horizontal frente al vertical. Los *porcentajes* serían, por tanto, *verticales* y las comparaciones horizontales.

Las diferencias porcentuales deben superar un determinado valor (al menos superior al 5%) para que puedan considerarse importantes. Depende del *error muestral* que derive de los tamaños de las *bases* sobre las que se calculan los *porcentajes*. Si estos *tamaños muestrales* son bajos, la diferencia porcentual ha de ser superior, si de ella quiere deducirse la existencia de *asociación* entre las *variables*.

En la exposición de la *tabla* (en el informe de la investigación) ha de indicarse, explícitamente, la dirección en la que se han calculado los *porcentajes*. Como sólo se aportan datos porcentuales, se recomienda poner entre paréntesis las *bases* de los *porcentajes*. Ello ayudará a la interpretación de las diferencias que en ellos se observen.

También se aconseja encabezar la *tabla* con un *título* que describa, sucintamente, el contenido de la *tabla*. En el *título* han de especificarse las *variables* comprendidas en la *tabla* y su relación.

A modo de ejemplo, la *tabla de contingencia* expuesta en el Cuadro 9.4 puede transformarse, en el *Informe,* como se expone en el Cuadro 9.5. Si bien, ha de matizarse que ésta constituye una de las posibles alternativas.

Pese a la existencia de convencionalismos en el formato de las tablas, el investigador es libre de diseñar el formato que más se ajuste a su estilo particular, y al contenido de la *tabla*.

La *tabla* expuesta en el Cuadro 9.5 constituye una *tabla simple* porque en ella figuran sólo dos *variables*. Pero también suelen componerse *tablas complejas,* a partir de

CUADRO 9.5. Menores clasificados por medida del tribunal, según el tipo de delito (porcentaje vertical.

Medida	Tipo de delito					
	Hurto	*Robo con intimidación*	*Robo sin intimidación*	*Insumisión paterna*	*Otros delitos*	*Total*
Amonestación	51	20	46	14	46	39
Libertad vigilada	33	59	37	59	40	42
Internamiento	16	21	17	27	14	19
Total	100 (78)	100 (61)	100 (253)	100 (73)	100 (63)	100 (528)

la conjunción de varias *variables* (independientes y/o dependientes). En estos casos, los análisis se centran en las variaciones en los *valores* extremos de las *variables*.

La *lectura porcentual,* aunque ilustrativa, resulta, no obstante, insuficiente. Precisa del complemento de *estadísticos* que gradúen la *asociación* entre las *variables* y su *significatividad*. Este complemento adquiere un mayor protagonismo cuando de las *tablas* quiera deducirse una *relación causal*.

Dos *variables* se hallan relacionadas si sus *atributos* varían conjuntamente. Para la graduación de esta relación, se acude a alguno de los *estadísticos de contingencia,* dependiendo del nivel de medición de la variable. Aquí sólo se enumeran. Consúltese en un manual de estadística su formulación.

a) Si la variable es *nominal,* los estadísticos que miden el grado de *asociación* entre dos *variables* son: Phi cuadrado, "C" de Pearson, "V" de Cramer, "Q" de Yule, Lambda, Tau-Y de Goodman y Kruskal...; además del coeficiente "d" o de diferencia de proporciones.
b) *Variables ordinales:* Rho de Spearman, Tau-A, Tau-B y Tau-C de Kendall, Gamma de Goodman y Kruskal, "D" de Sommer, entre otros.
c) *Variables de intervalo:* a los estadísticos anteriores se suma el coeficiente de correlación producto-momento de Pearson.

Cada uno de estos estadísticos indican la fuerza y la dirección de la *asociación* entre dos variables. Su *signo* expresa la dirección de la *correlación* (positiva o negativa); mientras que el *valor* numérico (que oscila entre 0 y 1), la magnitud de la relación. De este modo:

a) El valor ".00" denota *inexistencia* de *asociación*.
b) "–1.00", *correlación perfecta negativa* (conforme aumenta el valor de la variable independiente, disminuye el valor correspondiente a la dependiente).

c) "1.00", *correlación perfecta positiva* (al incremento de la variable independiente le sigue el aumento, también, de la dependiente).

Una vez medida la relación entre las variables, se comprueba su *significatividad*. Los datos analizados siempre pertenecen a una *muestra,* de las múltiples posibles, que pueden extraerse de una misma *población*. Razón por la cual, se exige la comprobación de la *significatividad* de los estadísticos y sus posibilidades de *inferencia* a la *población*.

En las *tablas de contingencia,* se aplica el *test de la Chi-Cuadrado* (X^2). Este estadístico se obtiene de la comparación entre las *frecuencias observadas* (en la *muestra)* y aquellas que cabría esperar en caso de inexistencia de relación entre las variables. Su valor se compara con el *teórico* (aquel que figura en una tabla de la X^2), para unos grados de libertad determinados y un nivel de significación escogido por el investigador. El *nivel de significación* habitual es .05. Este supone una posibilidad de error en la estimación del 5%. Los *grados de libertad,* en cambio, vienen marcados por las dimensiones de la *tabla:* número de *filas* (i) y de *columnas* (j). Concretamente, g. l. = (i–1)(j–1).

Cuando el valor de la X^2 empírico (el obtenido en la *muestra)* supera al teórico (el marcado en una tabla de la X^2), se deduce la *significatividad* estadística de la relación observada entre las *variables*. En caso contrario (X^2 empírico < X^2 teórico), se desestima la relación bivariable, por su no *significatividad*. La relación se consideraría, entonces, meramente casual, debida a *errores muestrales*.

En la salida de ordenador (véase Cuadro 9.4, por ejemplo) el valor de la X^2 aparece acompañado con su *significatividad ("significance")*. El valor que figura bajo este rótulo ha de ser inferior a .05 para que el valor de la X^2 sea significativo a un nivel de .05.

En cualquier manual de estadística en las ciencias sociales (como el de Blalock, 1978; García Ferrando, 1985; o Siegel, 1985) puede encontrarse una exposición detallada de estos y demás *estadísticos de contingencia*.

9.3.2. Otros análisis bivariables

Aparte de las *tablas de contingencia,* existen otras *técnicas de análisis bivariables,* como el de *regresión* y *varianza simple*. Ambas técnicas analíticas miden la relación de *dependencia* entre dos variables, si bien imponen mayores restricciones que las *tablas de contingencia*. Su cumplimiento exige que la *variable dependiente* sea *métrica* o cuantitativa. Ello determina su menor aplicabilidad en la investigación social, en la que predominan las *variables cualitativas (no métricas)*.

- *Varianza simple*

Esta técnica analítica es muy aplicada en los *diseños experimentales,* en la comprobación de los efectos de los tratamientos experimentales. A ello contribuye la es-

pecificidad del análisis. Su finalidad es comprobar la existencia de diferencias grupales respecto a una única *variable dependiente (métrica)*. Para ello se manipula una *variable independiente,* en función de cuyos valores se forman distintos grupos de tratamiento. Constituidos los grupos, se comprueba la media de cada uno de ellos respecto a la *variable dependiente*. Si se observan diferencias entre las *medias grupales,* se procede después a la comparación de las *varianzas grupales,* y a la medición de su *significatividad*.

Interesa que la *varianza* entre los grupos supere a la *varianza intragrupal*. Ello expresaría una mayor heterogeneidad entre los grupos, frente a una escasa variabilidad dentro de ellos. Por lo que podría afirmarse la existencia de diferencias entre los grupos.

La *significatividad* de las diferencias grupales se comprueba mediante los estadísticos "t" (si únicamente se han formado dos grupos de tratamiento), y "F" (si son más de dos los grupos creados).

En ésta, como en cualquier prueba de *significatividad*, se comparan los *valores empíricos* ("t" y "F") con los *teóricos* (mostrados en las tablas de la *"t" de Student* y de la *"F" de Fisher*, correspondientes). El proceso es similar al descrito en el *test de la Chi-Cuadrado*. Se fija el *nivel de significatividad,* en función de la precisión que el investigador desee para su estimación (.05, generalmente); y los *grados de libertad* (ahora determinados por el *tamaño muestral* y el número de *variables independientes* consideradas). Siempre que el *valor empírico* supere al *teórico,* las diferencias grupales observadas en la *muestra* adquirirán *significatividad estadística*. Podrán, por tanto, hacerse extensibles al *universo* del que se extrajo la *muestra* (en los niveles de probabilidad fijados).

• *Regresión simple*

Constituye otra técnica de *dependencia* en la que se analiza la relación entre una única variable *independiente (métrica o no métrica)* y una *dependiente (métrica)*. Pero, a diferencia de la técnica analítica anterior, la finalidad del análisis es la *predicción* del valor de la variable *dependiente* a partir del conocimiento de la *independiente*. Se cuantifica la relación existente entre ambas variables; y, se establece el grado de confianza o *significatividad* de la estimación efectuada.

La *correlación* entre las dos variables *(dependiente* e *independiente)* se mide mediante el *coeficiente R de Pearson*. Éste expresa el grado de *covariación* entre las variables, según se aproxime a "0" (inexistencia de asociación) o a "1" (asociación perfecta). También informa de la dirección de la asociación: creciente (si el signo es positivo) o decreciente (si es negativo).

En el *análisis de regresión,* la idea que subyace es la consecución de una *recta de regresión* que presente el mejor "ajuste" de los casos respecto a las variables analizadas. Esta *recta* tiene su expresión matemática en la siguiente *ecuación de regresión:*

$$y = a + bx + e$$

donde: "y" denota el valor de la variable *dependiente.*
"a" es el *intercepto* o punto de la *recta* que corta al eje de las Y.
"b" es la *pendiente* de la *recta* (también referido como el *coeficiente de regresión). Su *valor* expresa la cantidad de variación de la variable *dependiente* por cada unidad de variación de la *independiente.* Su *signo* denota si se produce aumento (pendiente creciente; signo positivo) o disminución (pendiente decreciente; signo negativo).
"e" representa el *error de la estimación:* la inadecuación de la *ecuación de regresión* en la predicción del valor de la variable *dependiente.*

Esta *ecuación* permite la predicción del valor de la variable *dependiente* a partir de valores conocidos de la *independiente.* Los *coeficientes* se obtienen, generalmente, siguiendo el *criterio de mínimos cuadrados* (hacer mínima la distancia que separa los *puntos* –obtenidos de la confluencia de ambas variables en cada uno de los casos– y la *recta de regresión).*

La *significatividad* de los *coeficientes* se comprueba mediante el estadístico "*t*", con n-1 *grados de libertad* (siendo *"n"* el número de observaciones). En cambio, la *significatividad* de la *correlación* se comprueba mediante el estadístico "*F*". Como en cualquier prueba de *significatividad,* los valores de "*t*" y de "*F*" *empíricos* han de superar los *teóricos* (determinados en las tablas respectivas) para que el modelo de *regresión* sea significativo estadísticamente.

9.4. El análisis multivariable

Los análisis *univariables* y *bivariables* con frecuencia se muestran insuficientes para cubrir los objetivos de la investigación. El proporcionar una visión conjunta e integrada, que describa y/o explique la realidad que se analiza, demanda la realización de *análisis multivariables* (de más de dos variables al mismo tiempo). De otra forma no podrían medirse las influencias e interrelaciones existentes entre grupos de variables. Como hace tiempo reconociera García Ferrando (1979: 198):

> "Las distribuciones bivariables en sociología aparecen demasiado simplistas para lograr adecuadas explicaciones científicas."

La peculiaridad del *análisis multivariable* reside en operar con un número elevado de variables, y de manera simultánea, basándose en el cálculo matricial. Kendall (1975) lo define como el conjunto de técnicas estadísticas que permite el análisis simultáneo de más de dos variables en una muestra de observaciones. A esta definición, Dillon y Goldstein (1984) añaden la posibilidad de análisis sincrónicos de mediciones en más de una *muestra.*

Para el *análisis multivariable* existe un amplio abanico de *técnicas*. En conformidad con un extenso grupo de autores (véase Kendall, 1975; Dillon y Goldstein, 1984; o Hair *et al.* 1992, por ejemplo), estas técnicas pueden agruparse en dos grandes categorías *(técnicas de dependencia y técnicas de interdependencia),* en función de si se diferencia, o no, entre variables *dependientes* e *independientes*.

En la elección de la *técnica* concreta a aplicar intervienen, básicamente, el objetivo de la investigación, y las características de las variables que se analicen (su número y nivel de medición).

A continuación se ofrece una breve descripción de cada una de las *técnicas multivariables* comúnmente referidas. Para un mayor conocimiento de cada una de ellas remito a la bibliografía especializada. Lo que sigue es una mera visión panorámica de las *técnicas multivariables*.

9.4.1. Técnicas multivariables de dependencia

Un conjunto de técnicas analíticas unidas por un mismo propósito: medir la existencia de *relaciones causales* entre un conjunto de variables, el grado y *significatividad* de la misma. Sin embargo, difieren en el número de variables *dependientes* que incluyen, y en el nivel de *medición* exigido *(métrico o no métrico)*. En la Figura 9.1 se esquematizan las principales alternativas, comúnmente señaladas, en el análisis de la *dependencia*.

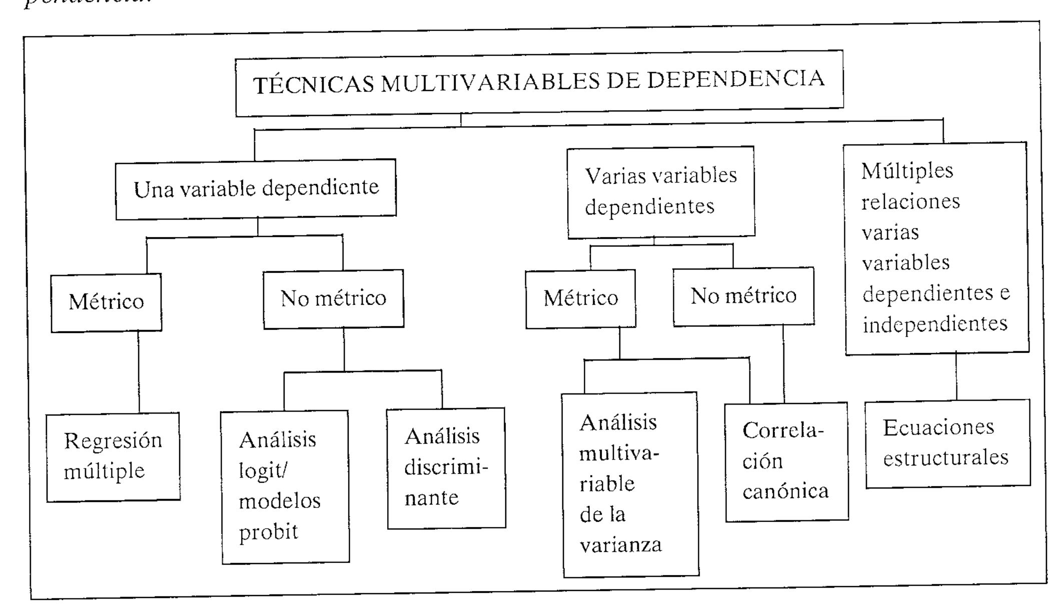


Figura 9.1. Técnicas de análisis multivariable de dependencia.

• *Regresión múltiple*

Una técnica de *dependencia* muy versátil, y utilizada, cuando se busca la *predicción* del valor de una variable *dependiente* (o *criterio),* a partir de valores conocidos en una serie de variables *independientes* (o *predictoras).* Para su realización se exige el cumplimiento de una serie de *supuestos* básicos:

a) *Linealidad:* la interdependencia entre las variables ha de responder a un modelo lineal. Esto quiere decir, que a cada cambio en una unidad en una variable corresponda un cambio, también en una unidad, en la otra variable.
b) *Normalidad:* la correspondencia de la distribución de los datos (para la variable *dependiente* y las *independientes)* con la *curva normal.* Esto permite la utilización de los estadísticos *"F"* y *"t"* para la comprobación de la *significatividad.*
c) *Homocedasticidad* o igualdad de las *varianzas* de los términos de error en la serie de variables *independientes.*
   Para que pueda medirse la relación de las variables *independientes* con la *dependiente* se precisa que la *varianza* de los valores de la variable *dependiente* sea igual en cada valor de las variables *predictoras.*
d) *Aditividad:* los efectos de las variables *independientes* han de poderse sumar entre sí, para poder predecir la variable *dependiente.*
e) Ausencia de *colinealidad* (de correlación) entre las variables *independientes,* con objeto de que puedan medirse sus efectos concretos en la variable *dependiente.*
f) Ha de haber un número elevado de observaciones. Al menos, debería haber 20 veces más casos que variables *independientes.* Máxime si el modelo de regresión se obtiene "paso a paso": incorporando una a una las variables *independientes* hasta que no exista ninguna más con poder predictivo significativo.
   Cuando no se obtiene esta proporción, habría que optar por alguno de los siguientes remedios: eliminar alguna variable *independiente,* o agrupar varias variables creando una nueva variable (que sea una combinación de dos o más variables *independientes).*

Como en *regresión simple,* en *regresión múltiple* se obtiene una *ecuación de regresión,* con tantos *coeficientes* como variables *independientes.* A partir de ella, podrá predecirse el valor medio de la variable *dependiente.* Asimismo, se mide el grado de *correlación* existente entre las variables (mediante el estadístico *R múltiple)* y su *significatividad* (gracias al estadístico *"F").*

El valor del R^2 *múltiple* (el *coeficiente de determinación)* expresará el porcentaje de *varianza* de la variable *dependiente* que es explicado por las *independientes.* Cuanto más elevado sea su valor, mejor para la predicción de la variable *dependiente.*

• *Análisis logit*

Los análisis *logit* y *probit* se definen como modelos analíticos alternativos (al modelo de *probabilidad lineal)* para variables *dependientes cualitativas.* Ambos modelos *(logit* y *probit)* se asemejan bastante. Si bien, el *análisis logit* ha alcanzado un mayor desarrollo, en parte debido a su mayor practicabilidad. De ahí que se le conceda un mayor protagonismo en esta breve reseña analítica.

El *análisis logit* constituye un modelo de respuesta cualitativa, que mide la relación entre una serie de variables *independientes (métricas* o *no métricas)* y una única variable *dependiente categórica (no métrica).* Para ello se basa en el análisis de la *razón de probabilidad* de una variable *dependiente,* en función de las *independientes.* De lo que se trata es de graduar la *probabilidad* de que determinadas variables *independientes* provoquen la ocurrencia de un evento concreto (la variable *dependiente).*

Del análisis se obtiene una *ecuación* similar a la de *regresión.* Los *coeficientes* (calculados siguiendo el criterio de *máxima verosimilitud,* o el de *mínimos cuadrados)* se interpretan como en *regresión.* La *significatividad* de cada uno de ellos también se comprueba mediante el estadístico *"t".*

La peculiaridad del análisis reside, no obstante, en el cálculo de los *incrementos de probabilidad.* Éstos se calculan para cada variable con *coeficiente* significativo. Se considera el valor del *coeficiente* (ß) y el *valor medio* de la variable en la *muestra* (X_j), siguiendo la expresión de McFadden (1974):

$$\mathrm{Prob}(Y=1)=F(X'_J\,\beta)=\frac{e^{X'_j\beta}}{1+e^{X'_j\beta}}$$

La *significatividad* del modelo global se comprueba mediante el estadístico X^2. Este estadístico se complementa con el *porcentaje de aciertos* (el porcentaje de observaciones que logran ser correctamente estimadas mediante los *coeficientes logit),* como un indicador del éxito del modelo en la predicción de la variable *dependiente.*

• *Análisis discriminante*

Una técnica de clasificación y de asignación de individuos a grupos, a partir del conocimiento previo de sus características.

Su objetivo básico es la estimación de la relación existente entre una serie de variables *independientes* y una única variable *dependiente no métrica* (o *categórica).*

El número de *categorías* que incluya la variable *dependiente* determina los grupos formados a partir de la *muestra.* Estos *grupos* deben, previamente, haberse configurado mediante la aplicación de una *técnica multivariable de interdependencia* (como el análisis de *conglomerados* o el *factorial).*

Con el *análisis discriminante* lo que se pretende es la comprobación de si los *grupos* creados por otra técnica analítica son adecuadamente caracterizados por las variables que les definen (de acuerdo con los análisis previos).

Asimismo, se trata de conocer cuál es la combinación de variables *(funciones discriminantes)* que hace máxima la diferenciación entre los grupos. El conocimiento de estas variables ayudará a la predicción de la probabilidad de pertenencia de un individuo concreto a uno de los grupos diferenciados. Ello dependerá de los valores que presente en las variables *independientes* analizadas.

La ejecución del *análisis discriminante* exige, igualmente, el cumplimiento de unos *supuestos* claves, como son:

a) La *normalidad* de las variables *independientes*. Se recomienda que éstas sean *métricas*. En caso contrario, habría que introducir modificaciones en el análisis.
b) *Matrices de covarianzas* iguales en cada grupo.
c) Inexistencia de *colinealidad* entre las variables *independientes,* para evitar que su información sea redundante en la explicación de la variable *dependiente.*
d) Ha de haber, al menos, 20 casos por cada variable *independiente* introducida en el análisis. Esto contribuye a la *significatividad* estadística del modelo obtenido.

La distancia entre los grupos se mide mediante el estadístico D^2 *de Mahalanobis*. A éste se añaden los estadísticos *lambda de Wilks, F* y X^2, en la comprobación de la *significatividad* de las variables *independientes* en la diferenciación entre los grupos.

La contribución de estas variables se cuantifica mediante los *coeficientes de función discriminante estandarizados* (similares a los *coeficientes beta* en *regresión)* y los *factores de carga discriminantes* (los *"discriminant loadings").* Sobre todo, por estos últimos, debido a su mayor precisión.

Como en el *análisis logit,* en el *discriminante* también se calcula la *razón de aciertos*, como medida del éxito del modelo en la predicción de la variable *dependiente* (la clasificación de los grupos).

• *Análisis multivariable de la varianza*

También conocido como MANOVA. Representa una extensión del análisis univariado de la varianza (ANOVA), mediante el cual se exploran, simultáneamente, las posibles relaciones existentes entre: varias variables *independientes no métricas* (normalmente referidas como *tratamientos)* y dos o más variables *dependientes métricas.* Lo que permite la medición de las *correlaciones* entre las variables *dependientes* y entre las *independientes*.

El uso de este análisis se adecúa, igualmente, a los *diseños experimentales*, en la comprobación de los efectos de distintos *tratamientos*. El fin que se pretende es la determinación de la existencia de diferencias, en los *valores medios* de las variables *dependientes,* en cada *grupo de tratamiento*. Ello exige que:

a) Las variables *dependientes* se hallen distribuidas *normalmente*.
b) Se parta de *matrices de varianza-covarianza* iguales en todos los *grupos*. En caso contrario, no podría medirse el efecto específico del *tratamiento* dado (debido a que se partiría de grupos inicialmente diferentes).
c) El *tamaño muestral* ha de superar al necesario para la *varianza simple,* si se pretende que el modelo adquiera *significatividad* estadística.

La *significatividad* de las diferencias multivariables entre los *grupos* se comprueba mediante cuatro criterios primordiales: la *raíz máxima de Roy,* la *Lambda de Wilks*, la *traza de Hotelling* y la de *Pillai*. Si estos tests multivariados resultan significativos, se procede a la comprobación de las diferencias grupales en cada una de las variables *dependientes*. Para ello se acude al estadístico "*F*".

• *Correlación canónica*

Una *técnica de dependencia* que permite la comprobación de la existencia de interrelación entre una serie de variables *dependientes* y otra serie de variables *independientes*. Ambas pueden ser tanto *métricas* como *no métricas.*

La finalidad principal del análisis es la obtención de *combinaciones lineales* de cada serie de variables *(dependientes* e *independientes).* Estas combinaciones han de maximizar las *correlaciones* entre las variables.

El grado de relación entre la serie de *variables canónicas* se mide mediante el *coeficiente de correlación canónica.* Éste, elevado al cuadrado (R^2), representa el porcentaje de *varianza* de una combinación de variables *dependientes canónicas* que es explicada por una combinación de las variables *independientes*. Su *significatividad* se comprueba, de nuevo, mediante el estadístico *"F"*.

Las contribuciones de cada variable (*dependiente* e *independiente*) a la combinación o serie de *variables canónicas* respectiva se comprueba, al igual que en el *análisis factorial* y *discriminante,* mediante los *factores de carga ("canonical loadings").* Estos han de ser ≥.30 para que se consideren relevantes. La contribución de las variables también puede comprobarse mediante los *pesos canónicos ("canonical weights"),* si bien estos últimos presentan una mayor inestabilidad que los anteriores (los *factores de carga).*

• *Ecuaciones estructurales*

También conocidas como modelos LISREL, en referencia a uno de los programas estadísticos más populares para su resolución.

Esta última *técnica de dependencia* puede catalogarse como una extensión del *análisis factorial* y de *regresión múltiple,* por dos razones fundamentales:

a) En las relaciones de *dependencia* se representan variables *latentes* o constructos (no observadas), a partir de valores conocidos en las variables *manifiestas* (u observadas) –a semejanza con el *análisis factorial*–.

b) Las diversas *relaciones causales* tienen su expresión matemática en varias *ecuaciones de regresión* (ahora denominadas *ecuaciones lineales estructurales).*

Pero, a diferencia de las *técnicas de dependencia* anteriormente expuestas, en los *modelos de ecuaciones estructurales* se comprueba (simultáneamente) diversas *relaciones causales.* No sólo entre varias variables *independientes (métricas o no métricas)* y *dependientes (métricas),* sino entre ellas mismas, también (dependientes con dependientes e independientes con independientes). De ello resulta un *modelo causal* de mayor complejidad, aunque más ajustado a la pluridimensionalidad de las *relaciones causales.*

Las *relaciones causales* también se representan gráficamente, mediante un diagrama de "sendero" *("path").* En él se diferencian las variables *latentes* de las *observadas,* además de los *errores de medición* (de cada modalidad de variable), y el tipo de relación entre ellas (por ejemplo, una flecha bidireccional indica la existencia de correlación entre dos variables).

Para cada variable que recibe una flecha se formula una *ecuación.* En esa *ecuación,* dicha variable actúa como variable *dependiente.*

La realización de esta técnica analítica exige el cumplimiento de los mismos supuestos que en las anteriores (como la *independencia* de las observaciones; la *linealidad* de las relaciones; la *normalidad* multivariable; la selección *aleatoria* de las unidades muestrales; y un elevado *tamaño muestral).*

Los *parámetros* se obtienen mediante el método de *mínimos cuadrados generalizados,* o por el de *máxima verosimilitud.* Si bien, el primer método muestra imprecisión, conforme desciende el tamaño de la *muestra* y aumenta la complejidad del modelo. La *significatividad* de los *parámetros* se comprueba mediante la *"t",* además de considerarse el *error* de la estimación.

Como se diferencia entre variables *latentes* y *observadas,* se comprueba la *fiabilidad* (≥.70) y la *validez* (≥.50) de la *medición* de los *constructos teóricos (variables latentes).* A ello se suma el *ajuste* global del modelo alcanzado respecto al inicial.

Existen distintos *índices de ajuste.* Entre ellos destacan los siguientes: el estadístico X^2 (aunque ahora interesan los valores bajos, porque expresan un mejor *ajuste* entre el modelo obtenido y el propuesto inicialmente); el *índice GFI,* que oscila entre .00 y 1.00 (cuanto más se aproxime a 1, mejor); los *índices TLI* (≥.90), *NFI* (≥.90), *AGFI* (≥.90), y *AIC* (≥.70), entre otros.

9.4.2. Técnicas multivariables de interdependencia

A diferencia de las técnicas analíticas anteriores, las de *interdependencia* presentan un menor poder predictivo. Mediante ellas se analiza la existencia de *asociación* o

relación mutua entre varias variables, sin diferenciar entre *dependientes* e *independientes*.

La Figura 9.2 representa las principales técnicas de interdependencia. La elección entre una u otra responde a los mismos criterios fundamentales mencionados en las *técnicas de dependencia*. Concretamente, el objetivo de la investigación y el nivel de medición de las variables. Ahora la diferencia básica entre las técnicas se establece en función del nivel de *medición* mínimo exigido en las variables para su cumplimentación: *métrico* o *no métrico*.

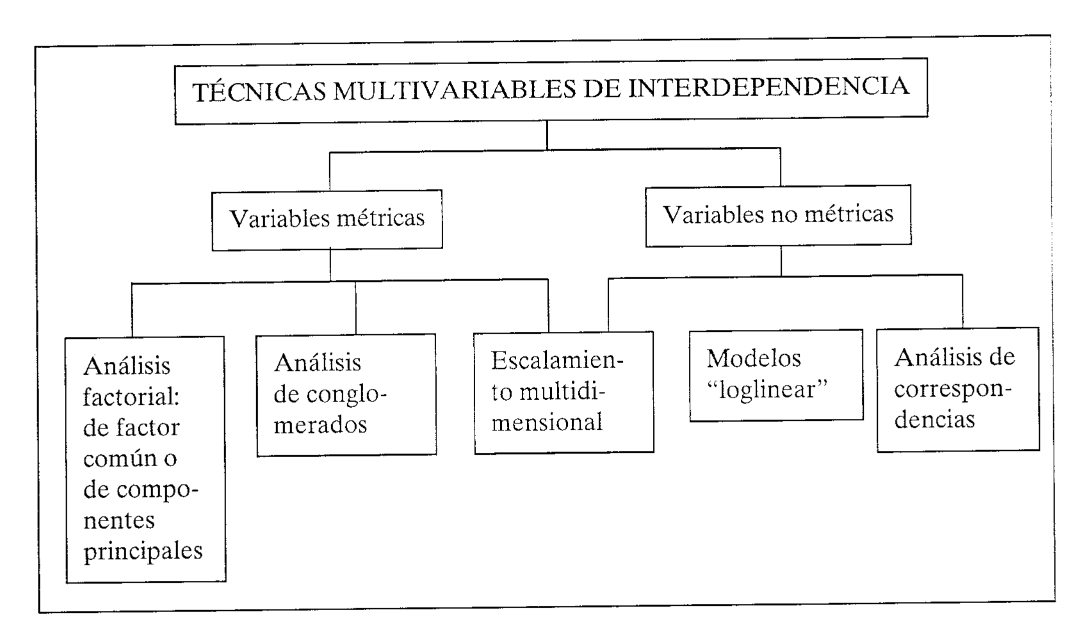


Figura 9.2. Técnicas de análisis multivariable de interdependencia.

• *Análisis factorial*

Una denominación genérica que engloba distintos procedimientos centrados en el análisis de la variación total que una variable comparte con otras variables. Su objetivo fundamental es resumir, la información contenida en un conjunto de variables interrelacionadas, en un número reducido de *dimensiones latentes* comunes (o *factores)*.

El *análisis factorial* es *exploratorio,* cuando no se conoce, previamente, cuáles son los *"factores"*. Éstos se determinarán, precisamente, tras el *análisis factorial.* En cambio, el análisis será *confirmatorio,* si se parte de unos *"factores"* especificados *a priori,* mediante un análisis *factorial exploratorio* u otra técnica multivariable (como el análisis de *conglomerados).* Estos *factores* representan a un conjunto de variables *empíricas* u observadas. Entonces, se tratará de "corroborar" la adecuación de estas variables (empíricas) en la medición de las *dimensiones* de los *conceptos teóricos*.

Además de esta clasificación básica, se diferencia entre análisis factorial de *componentes principales* y de *factor común*. En el *análisis de componentes* se persigue la obtención de *combinaciones lineales* de variables que logren explicar la mayor proporción de *varianza* conjunta. Para ello se tiene en cuenta tanto la *varianza específica* de cada variable, como la *varianza conjunta* (o compartida con otras variables). Por el contrario, en el *análisis de factor común,* la identificación de los *factores latentes* responde sólo a la *varianza común* de una serie de variables.

En cualquiera de las modalidades de *análisis factorial,* el investigador deberá concretar el método para la extracción de los *factores: ortogonal* (quartimax, varimax o equimax) u *oblicuo*. En el *ortogonal,* los *factores* se extraen de manera que sean totalmente independientes unos de otros. La extracción *oblicua* permite, en cambio, la existencia de correlación entre los *factores*.

En la decisión de cuántos *factores* escoger han de compaginarse distintos criterios como, por ejemplo, el criterio de *raíz latente (autovalor* superior a 1), o el de *porcentaje acumulado de varianza* (al menos superior al 60%).

Las *correlaciones* de las variables *empíricas* con los *factores* se comprueban mediante los *factores de carga ("factor loadings")*. Éstos han de ser ≥.30 para considerarse significativos.

Una vez definidos los *factores,* se les asigna un nombre o *etiqueta*, que refleje el contenido de las variables *empíricas* que representa.

Las investigaciones comentadas en los Capítulos 4 y 10 ilustran la aplicación de esta técnica analítica. Razón por la que se remite a su lectura.

• *Análisis de conglomerados*

Una variedad de análisis cuya finalidad principal es la clasificación de un grupo de individuos u objetos en un número reducido de grupos. Estos *grupos* han de ser mutuamente excluyentes; han de estar compuestos por individuos lo más similares posible entre sí y diferentes de los integrantes de otros grupos.

A diferencia del *análisis discriminante,* los *grupos* no se hallan predefinidos. Precisamente se aplica esta técnica analítica para la definición de los grupos. Posteriormente, éstos pueden ser corroborados mediante otras técnicas multivariables (como el análisis *discriminante* o el *factorial confirmatorio)*.

El *grupo* se define en función del algoritmo de clasificación que se emplee en la agrupación de los sujetos. En general, se busca la agrupación que haga máxima la distancia entre las *medias grupales* y mínima la *desviación intragrupal* (de los individuos integrantes del grupo respecto a su *centroide* o *media grupal)*.

Para la obtención de estos grupos, o *conglomerados,* puede elegirse entre distintos procedimientos, agrupados en dos amplias categorías: *jerárquicos* y *no jerárquicos*. Los primeros se dividen, a su vez, en *aglomerativos* y *disociativos* (o *divisorios)*.

a) El procedimiento *jerárquico aglomerativo* comienza con tantos grupos como individuos en la *muestra*. Paulatinamente van reduciéndose el número de grupos

y aumentando, en contra, el número de individuos en cada grupo. El programa finaliza cuando un único *conglomerado* agrupa a todos los individuos.

b) El procedimiento *jerárquico disociativo* (o *divisorio)* procede a la inversa. Parte de un único grupo, que integra a todos los individuos. Poco a poco, este grupo va seccionándose en diferentes subgrupos hasta que, al final, existen tantos grupos como individuos.

En ambos procedimientos *jerárquicos* de formación grupal, la elección del número de *conglomerados* se realiza conforme a una variedad de criterios. Entre ellos destacan: el criterio de la *distancia mínima;* de la *distancia máxima;* el *promedio de las distancias;* el *método Ward;* o el de los *centroides.* Mientras que los primeros criterios consideran la distancia de los individuos, el último tiene en cuenta la distancia entre las *medias (centroides)* de las variables. En cualquiera de estos procedimientos *jerárquicos* los datos pueden visualizarse mediante un *dendograma:* una representación gráfica en forma de árbol.

En los procedimientos de agrupación *no jerárquica* se parte, a diferencia de los anteriores, de una especificación previa de los grupos que desean formarse. De lo que se trata es de encontrar representantes para cada uno de los grupos. Un individuo pertenecerá al grupo cuya distancia a su *centroide* sea menor.

Una vez que los grupos se han constituido (por cualquiera de los procedimientos referidos), se procede a la comparación de las *varianzas* de las variables en relación a los grupos *(intergrupo* e *intragrupo).* Ello permite comprobar si los grupos presentan diferencias en los valores de las variables consideradas en el análisis.

En función del valor del estadístico F (en cada una de las variables), se rechazan aquellas variables que no diferencian a los grupos. A partir del valor de la *media* de las variables se extrae, por último, los *rangos* que más separan a los integrantes de cada grupo.

• *Escalamiento multidimensional*

Una variedad analítica análoga al *análisis factorial.* Con él comparte un mismo objetivo: la obtención de un número reducido de *dimensiones,* que permitan caracterizar a determinados objetos o sujetos. Si bien difiere (entre otros aspectos) en el número de *dimensiones* a obtener. Mientras que el *análisis factorial* no impone restricciones al respecto (de hecho el análisis puede efectuarse con un número elevado de *dimensiones* o *factores),* el *escalamiento multidimensional* aconseja su reducción al menor número posible. Ello responde a condicionamientos impuestos para la representación gráfica de los resultados de la investigación.

En el *escalamiento multidimensional* los datos se representan como puntos en un espacio "multidimensional". La distancia habida entre ellos se considera una expresión gráfica de su semejanza o disimilaridad. Esta representación se hace más fac-

tible e interpretable cuando la información puede reducirse a dos o tres *dimensiones,* como máximo.

Para la concreción de estas *dimensiones* se acude a uno (o varios) *criterios de bondad de ajuste:* el "stress" de Kruskal, el "s-stress" de Young, la *correlación múltiple al cuadrado,* o el *diagrama de Shepard.*

Para que la configuración de las *dimensiones* presente un *ajuste* adecuado, los valores correspondientes a "stress" y "s-stress" han de ser bajos. En cambio, la *correlación múltiple al cuadrado* ha de ser elevada. Su valor se interpreta como proporción de *varianza* explicada por las distancias respectivas. Por su parte, el *diagrama de Shepard* ha de reflejar una tendencia ascendente (o creciente).

Dependiendo del nivel de medición de las variables, se diferencia entre *escalamiento multidimensional "métrico"* y *"no métrico".*

a) *Métrico,* cuando las variables son de *intervalo* o de *razón.* Lo que favorece su configuración en una escala *continua.*

   A partir de una *matriz de correlaciones,* o de *distancias* entre objetos, se trata de situar a éstos en un espacio *multidimensional.* Para ello se transforma la similaridad o disimilaridad percibida en ellos, en distancias *(euclidianas).*

b) *No métrico,* si las variables son *cualitativas (nominales* u *ordinales).* En este caso, se parte de una *matriz de rangos.* De ésta se obtiene información de la similaridad de los objetos. La configuración final será aquella que mejor represente a los *rangos* de la *matriz* inicial.

Como en el *análisis factorial,* el *escalamiento multidimensional* puede ser *exploratorio* y/o *confirmatorio.* Depende de la finalidad del mismo.

• *Modelos "log-linear"*

También conocidos como modelos "lineales logarítmicos" o modelos "log-lineales". Su denominación deriva de la transformación logarítmica operada en los datos (las frecuencias observadas en las variables) para facilitar su aditividad: la posibilidad de sumar los distintos efectos de un conjunto de variables *no métricas* (a semejanza de las variables *métricas* en el *análisis de regresión).*

Las *tablas de contingencia* resultan muy útiles en el análisis de variables *cualitativas.* Pero, cuando se añaden terceras o cuartas variables, a modo de variables de *control,* la interpretación de la *tabla* se complica. En estos casos, sobre todo conforme aumenta el número de variables, se aconseja la aplicación de *modelos log-linear.* Éstos se muestran adecuados para el análisis de las interrelaciones entre una serie de variables *no métricas,* que conforman una *tabla de contingencia multidimensional.*

Las casillas de la *tabla* se traducen a componentes, denominados *parámetros lambda.* Estos miden el efecto de los valores de las variables que conforman cada casilla. Si

el valor Z correspondiente a cada *lambda* es ≥ 1.96, el *parámetro lambda* se considera relevante en la interpretación del modelo. De él se pretende obtener la probabilidad de que un individuo concreto comparta una combinación específica de atributos de un conjunto de variables.

Para la comprobación del *ajuste del modelo* se acude a los estadísticos X^2 *y la razón de verosimilitud*. Los valores de ambos estadísticos tienden a coincidir conforme aumenta el tamaño de la muestra.

• *Análisis de correspondencias*

Una de las *técnicas de interdependencia* de más reciente desarrollo, también adecuada al análisis de variables *cualitativas*.

Parte de la configuración de las variables en una *tabla de contingencia*. Su objetivo fundamental es la representación de las distancias de las *filas* y las *columnas,* que integran la *tabla,* en unos *ejes cartesianos*. Para ello se transforma el valor de la X^2 en una medida métrica de distancia.

Como en el *escalamiento multidimensional,* la proximidad de los puntos mide la similaridad existente entre ellos. En el *análisis de correspondencias,* la proximidad muestra, concretamente, la asociación entre las categorías de las variables.

Primero se procede (como en el *análisis factorial* y el *escalamiento multidimensional)* a la identificación del número adecuado de *dimensiones* que categoricen al objeto de estudio. A tal fin, se examina, igualmente, el porcentaje acumulado de *varianza* explicada.

Para cada *dimensión* se deriva, a continuación, unos *autovalores ("eigenvalues")*. Éstos expresan la contribución relativa de cada *dimensión* en la explicación de las *varianzas* de las variables.

Una vez establecida la *dimensionalidad,* se identifica la *asociación* o relación existente entre las categorías de las variables mediante su *proximidad*. Ésta ha de comprobarse o en las *filas* o en las *columnas*. Depende de su ubicación. Un valor elevado de X^2 indica un fuerte grado de "correspondencia" entre los atributos de las variables.

De esta forma se obtiene la reducción dimensional de las proporciones de objetos en una serie de atributos. Al mismo tiempo, se extrae la representación de los objetos, relacionados con esos atributos.

9.5. Paquetes estadísticos disponibles

Sin la mediación del ordenador no sería viable ninguno de los análisis estadísticos reseñados. Especialmente, los análisis *multivariables*.

La gran revolución experimentada en los últimos años en el campo de la informática hace que la información que se dé quede obsoleta en un breve período de tiem-

po. Por esta razón, únicamente van a nombrarse algunos de los principales *paquetes estadísticos,* para conocimiento del lector.

La distinción entre los *paquetes estadísticos* responde a su especificidad. Se diferencia entre programas "genéricos" (que ejecutan la generalidad de los análisis estadísticos), y los "específicos" (especializados en técnicas analíticas concretas).

Entre los *paquetes estadísticos genéricos* destacan el SPSS, BMDP, SAS, ESP, y OSIRIS. En sus manuales respectivos se detallan tanto aspectos técnicos (de funcionamiento del programa), como analíticos (para la interpretación de los análisis estadísticos). Dicha interpretación será la misma, independientemente del programa utilizado.

Los *paquetes estadísticos específicos* son numerosos y variados. A modo de ejemplo se señalan los siguientes: LISREL y EQS (para el análisis de *ecuaciones estructurales),* LIMDEP (análisis *logit),* INDSCAL *(escalamiento multidimensional),* ECTA *(tablas de contingencia),* SPAD (análisis de *correspondencias),* o BROCOLI *(series temporales).*

A estos paquetes estadísticos se suman otros específicos para la realización de *gráficos* (como el SYSTAT, STATGRAPHICS, o el HARVARD GRAPHICS). También hay que mencionar las *bases de datos numéricas* (como DBASE o SIR-DB). Estas *bases* están diseñadas para la exportación automática de datos en la mayoría de los paquetes estadísticos (en especial, los *genéricos).*

Finalmente, cabe mencionar que en los últimos años comienzan a comercializarse *sistemas expertos en análisis de datos* (como el GLIM). Éstos funcionan a modo de consultor experto en *técnicas cuantitativas de análisis.*

Lecturas complementarias

Bisquerra, R. (1989): *Introducción conceptual al análisis multivariante.* Barcelona, PPU.

Garrido Luque, A.; J. L. Álvaro Estramiana (1995): *Técnicas de análisis estadístico en ciencias sociales.* Madrid, Universidad Complutense.

Hair, J. *et al.* (1992): *Multivariate data analysis.* New York, McMillan.

Sánchez Carrión, J. J. (comp.) (1984): *Introducción a las técnicas de análisis multivariable aplicadas a las ciencias sociales.* Madrid, CIS.

Sánchez Carrión, J. J. (1989): *Análisis de tablas de contingencia: el uso de los porcentajes en las ciencias sociales.* Madrid, CIS, *Monografía* n.° 105.

Sánchez Carrión, J. J. (1995): *Manual de análisis de datos.* Madrid, Alianza Universidad.

Sánchez Carrión, J. J. y Torcal, M. (1992): *Utilidades del SPSS/PC+. Presentación de informes. grabación de datos y creación de gráficos y mapas.* Madrid, Alianza.

Spiegel, M. (1991): *Estadística. Teoría y problemas resueltos.* Madrid, McGraw Hill.

Ejercicios Propuestos

1. Interprete los datos que figuran en los Cuadros 9.3. y 9.4.
2. Describa una investigación en la que se haya aplicado alguna técnica analítica multivariable. Especifique los análisis efectuados y su adecuación respecto a los objetivos de la investigación.
3. Trace el plan de análisis que seguiría para obtener el perfil del alumnado de la facultad, añadiendo las razones de su elección.
4. Si la investigación tuviera como objetivo comprobar qué variables influyen en la probabilidad de aprobar una asignatura, ¿qué análisis efectuaría? Justifique la respuesta.
5. A partir de los datos siguientes, confecciones una tabla e interprétela.

Edad	*Sexo*	*Ideología política*	*Frecuencia absoluta*
De 18 a 25 años	Varón	Izquierda	130
		Centro	220
		Derecha	275
	Mujer	Izquierda	115
		Centro	105
		Derecha	310
De 26 a 40 años	Varón	Izquierda	175
		Centro	270
		Derecha	210
	Mujer	Izquierda	196
		Centro	307
		Derecha	150
De 41 a 60 años	Varón	Izquierda	320
		Centro	178
		Derecha	97
	Mujer	Izquierda	205
		Centro	240
		Derecha	170

10

EL ANÁLISIS DE CONTENIDO CUANTITATIVO

En el análisis de *documentos* (y, en general, de cualquiera de las modalidades de comunicación verbal, e incluso visual) adquiere un destacado protagonismo el *análisis de contenido*. Aunque algunos autores (Krippendorf, 1980, 1990; y algunos de sus seguidores: López Aranguren (1989), O'Connell y Layder (1994), entre otros) lo cataloguen como "técnica de investigación", considero que las características del *análisis de contenido* corresponden más bien a una "técnica de análisis", de utilidad en cualquiera de las *estrategias de investigación* anteriormente referidas. Desde el *uso de fuentes documentales* y el *estudio de casos,* hasta la propia *encuesta* (en la *codificación* de *preguntas abiertas).*

La posibilidad de utilización de esta técnica surge siempre que el análisis se centre en el "contenido" *manifiesto* y *latente* de la información "verbal" (extraída de distintas fuentes: documentos, prensa escrita, transcripciones de entrevistas, grabaciones radiofónicas, de programas de televisión) o "visual" (de la observación de imágenes estáticas –fotografías, cuadros– y/o en movimiento –una grabación en vídeo, por ejemplo).

Pero, ¿en qué consiste el *análisis de contenido?,* ¿qué peculiaridades distinguen a esta variedad de análisis?, ¿cuáles son sus límites y posibilidades? A estas y otras preguntas tratará de darse respuesta en las páginas siguientes.

10.1. El análisis de contenido como técnica de análisis

El *análisis de contenido* se perfiló, en sus orígenes, como una técnica analítica eminentemente *cuantitativa.* Berelson (1952) la definió como la descripción *objetiva, sistemática* y *cuantitativa* del *contenido manifiesto* de la comunicación. No obstante, en la actualidad se reconoce la pluralidad analítica existente, de hecho, bajo la rúbrica

común del "análisis de contenido". Éste no se limita a la "cuantificación" del *contenido manifiesto* de la comunicación; sino que también aborda la interpretación del *contenido latente*.

Tesch (1992) diferencia dos tipos genéricos de *análisis de contenido:* el *clásico* y el *etnográfico*. El primero se ajusta a la definición de Berelson (1952), mientras que el segundo (el *etnográfico),* responde a una redefinición desde posturas *cualitativas* (Altheide, 1987), que pone el acento en el *análisis reflexivo de documentos*. Este segundo tipo se dirige, no tanto a la *descripción,* como a la *comprensión* de los *significados latentes,* y a la *verificación* de relaciones teóricas.

Dada la especificidad del presente manual (sobre la *metodología cuantitativa),* la exposición que aquí se hace del *análisis de contenido* se ciñe a la modalidad *clásica* o *cuantitativa*. Para su complementación con el *análisis de contenido etnográfico,* remito al manual de metodología cualitativa de esta misma colección, y a cualquier texto especializado en el *análisis cualitativo*.

10.1.1. Características esenciales del análisis de contenido cuantitativo

Los tres rasgos definitorios del *análisis de contenido cuantitativo* están ya presentes en la definición clásica de Berelson (1952): *objetividad, sistematicidad* y *cuantificación* de los *contenidos manifiestos* de la comunicación. A éstos se suma otro fundamental: la *inferencia* de los "datos" al "contexto" de referencia. De hecho, éste es –siguiendo a Krippendorff (1980, 1990)– el propósito primordial del análisis: realizar, a partir de ciertos datos, *inferencias* válidas y replicables, que puedan aplicarse a su *contexto*. Ello exige que el análisis reúna condiciones mínimas de *validez* y de *fiabilidad,* en el proceso de medición.

En el *análisis de contenido cuantitativo* el énfasis no recae en los aspectos semánticos o sintácticos de los textos, sino en la "cuantificación" de sus integrantes (palabras, expresiones, frases, temas); es decir, en la medición de su "frecuencia" de aparición en el texto.

Esta "cuantificación" permite, precisamente, la comparación de los "contenidos" de diferentes documentos y, sobre todo, la *inferencia*. Ésta sería –de acuerdo con Weber (1994)– difícilmente alcanzable por otros medios. Pero, para que pueda, a partir de los *contenidos "manifiestos",* inferirse los *contenidos "latentes"* (no explícitos en el documento, y relativos tanto al mensaje, como a sus agentes emisores y receptores), la "cuantificación" ha de hacerse de forma *sistemática* y *objetiva*.

La *objetividad* se halla relacionada con el uso de procedimientos "rigurosos y replicables" (Scott, 1990: 130). No sólo se han de elegir *indicadores* que constituyan una representación "válida" de los *conceptos*. También ha de lograrse la coincidencia en la *codificación* de un mismo documento, por personas diferentes, si se pretende que la clasificación efectuada se considere "fiable".

En suma, el *análisis de contenido cuantitativo* puede caracterizarse como un análisis no limitado a la *descripción,* sino orientado a la *inferencia*. La *descripción* (o enu-

meración de las características del texto) constituye –en conformidad con Bardin (1986)– la primera etapa del análisis; la *interpretación* (la significación acordada de estas características), la última. Entre ambas se encuentra la *inferencia,* como un procedimiento intermedio, que permite el paso de la *descripción* a la *interpretación.*

10.1.2. Fases de un análisis de contenido

Bardin (1986) diferencia tres momentos cronológicos fundamentales en la organización del *análisis de contenido:*

a) El *preanálisis* o fase de organización, de *operacionalización* de las ideas originarias de la investigación. Incluye:
   1) La especificación de los *objetivos* de la investigación y la formulación de las *hipótesis.*
   2) Elaboración de los *indicadores* en que se apoyará la interpretación de los datos. Éstos se obtendrán a partir de las *definiciones operativas* de los *conceptos teóricos* básicos en la investigación.
   3) Elección de los *documentos* a analizar: la *muestra* del estudio.

b) Explotación del material. Incluye:
   1) Elección de las *unidades de codificación* (palabras, expresiones, frases, temas, imágenes...; depende de qué se analice). De ellos, se cuantificará la *frecuencia,* e *intensidad,* de aparición en el documento. También pueden efectuarse enumeraciones de carácter temporal (cuándo) y espacial (dónde aparecen).
   Además de las *unidades de codificación,* se identifican las *unidades de contexto,* siguiendo la recomendación de Holsti (1969). Una misma palabra, por ejemplo, puede adoptar significados diferentes, dependiendo del *contexto* donde se ubique. Razón por la que se aconseja su contextualización. Ésta ayudará en la *interpretación* del significado de las *unidades de codificación.*
   2) Determinación de las *categorías* a emplear en la *codificación* del "contenido" de los documentos, a partir de las *hipótesis* de la investigación.
   Esta fase es exigida, cuando se realiza un *análisis de contenido categorial* o temático. Si bien, ésta no es la única variedad de *análisis de contenido,* aunque sea la más habitual.
   3) El *registro* y *cuantificación* de los *datos,* de acuerdo con los *indicadores* elegidos.

c) Tratamiento e interpretación de los resultados. A los datos reunidos se les da tratamiento estadístico, a semejanza de lo descrito en el Capítulo 9. Se comienza con el *análisis univariable.* Si las características de los datos lo permiten. puede concluirse con la aplicación de alguna técnica de análisis multivariable.
   Los resultados se someten a *pruebas estadísticas* y a *tests de validez.* En caso afirmativo, el investigador podrá hacer *inferencias* e *interpretaciones* de los

resultados de la investigación. Estas *interpretaciones* pueden actuar como punto de partida en una nueva investigación.

En la Figura 10.1 se resumen las fases principales en un *análisis de contenido cuantitativo*.

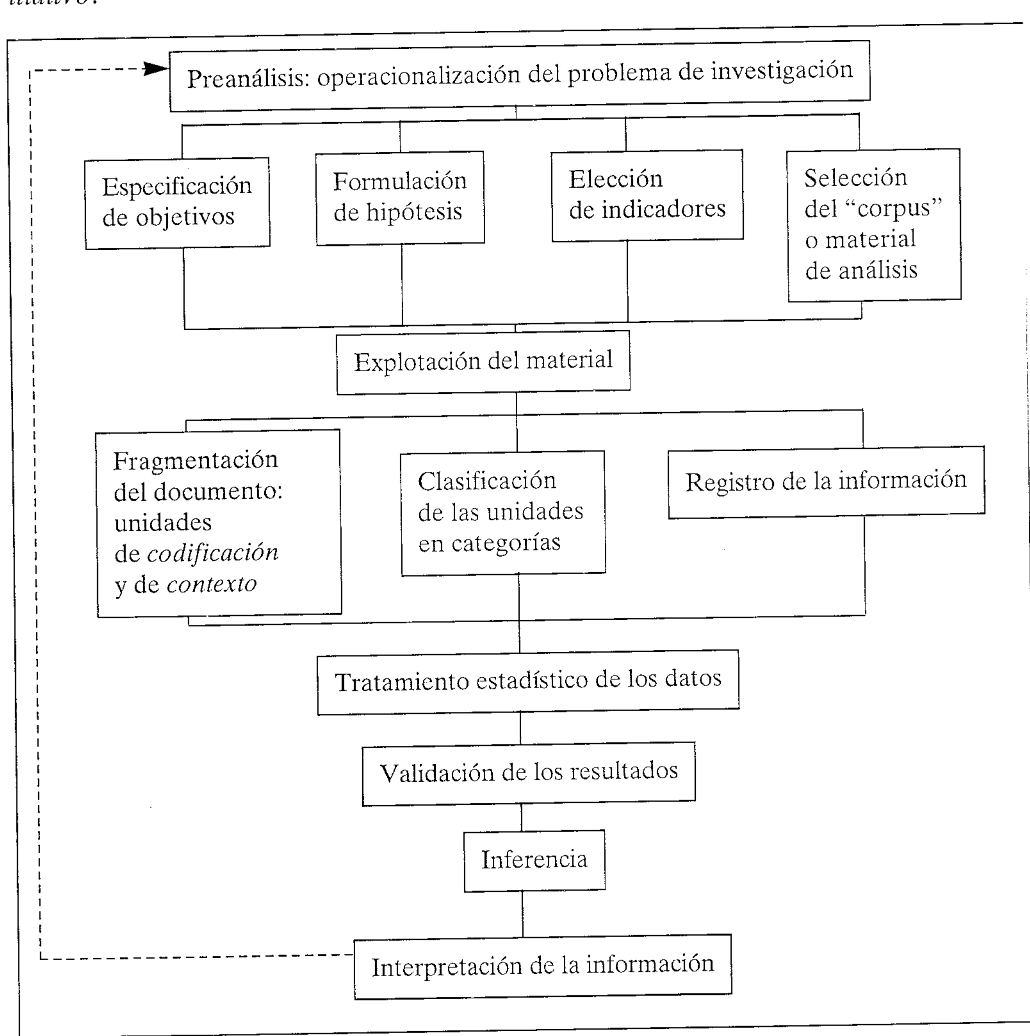


Figura 10.1. Fases principales en la relización de un análisis de contenido cuantitativo.

De las fases representadas en el gráfico, únicamente se comentarán, a continuación, los aspectos más específicos del *análisis de contenido*. Sobre los demás, remito a los capítulos iniciales del libro.

10.2. La selección del "corpus" de documentos o material de análisis

En función de los objetivos de la investigación, se eligen las *unidades de análisis* y la *estrategia* a seguir en la recogida de información.

Para efectuar un *análisis de contenido*, previamente ha de delimitarse el *"corpus"* o material a analizar. Esta delimitación incluye tanto a la *población* de documentos, como la concreción del procedimiento de *muestreo* a seguir en su selección.

Salvo que el *universo* sea de escasa amplitud, lo habitual es proceder a la extracción de una *muestra* representativa de dicho *universo*. En esta selección se siguen los mismos procedimientos muestrales descritos en el Capítulo 5. Concretamente, se consideran las características de la población de documentos (escritos, visuales, etc.), y el tipo de *inferencia* que desea realizarse. En general, se recomienda seguir procedimientos de selección *aleatorios*. Con estos se logra una mayor representatividad de la *muestra*.

La *selección muestral* comienza con la elección de la *fuente de comunicación* (periódicos, revistas, libros, canciones, imágenes). Prosigue con la extracción de *documentos* y, en caso de que éstos sean muy extensos, de *partes* de estos documentos.

EJEMPLO DE UNA SELECCIÓN MUESTRAL PARA UN ANÁLISIS DE CONTENIDO

Imaginemos que se desea analizar el tratamiento dado a los temas de corrupción en la prensa española. Uno de los posibles *diseños muestrales* incluiría las *fases* siguientes:

a) *Elección de fuentes.* Habría que decidir qué periódicos representarán el "*corpus*" de la investigación. El investigador puede optar entre elegirlos "intencionalmente", o seleccionarlos de forma *aleatoria*, siguiendo cualquiera de los procedimientos expuestos en el Capítulo 5 (desde el *muestreo aleatorio simple*, hasta el de *conglomerados).*

Si se busca una *muestra* representativa de la pluralidad de periódicos nacionales (de ámbito estatal, autonómico, o provincial), puede hacerse un *muestreo aleatorio estratificado* por Comunidades autónomas e ideología de las editoriales de los periódicos.

b) Seleccionados los periódicos, se procede a la *elección,* en cada uno de ellos, de unos *números* concretos. En esta segunda fase del *diseño muestral,* la selección atiende, habitualmente, a "fechas" de publicación. Como indican Sánchez Carrión (1985) y Weber (1985, 1994), los "contenidos" de los periódicos difieren por día de la semana (laboral o festivo) y período del año (diferenciando, sobre todo, el período vacacional del resto del año). Estos y otros aspectos deberán tenerse presentes en la selección (intencional o aleatoria) de las *fechas* de edición de los periódicos que forman la *muestra.*

c) *Extracción de textos en los documentos de la muestra,* si las dimensiones de éstos lo exigen. En este último estadio del *diseño muestral,* la selección suele ser, a diferencia de las anteriores, *intencional.* Se buscan, de forma intencional, "aquellas unidades que tratan sobre el tema objeto de estudio" (Sánchez Carrión, 1985: 96).

Weber (1994) recomienda el análisis del texto completo, siempre que sea posible. Ello ayudará a preservar la coherencia semántica del texto como unidad. Cuando nc sea factible el análisis del documento completo, el investigador deberá considerar la *estructura del texto* en la *selección muestral*. En concreto:

1) Podrían excluirse las secciones de *introducción* y de *cierre*.
2) Cada *muestra* debería constar de uno o más *párrafos* enteros para no alterar la coherencia semántica.
3) Las *frases* no deberían muestrearse. Analizar frases separadas (incluso aquellas que se han extraído de un mismo tema), destruye la coherencia semántica. Esto dificulta la *validación* posterior e *interpretación* de los resultados del *análisis de contenido*.

Por último, respecto al *tamaño de la muestra,* éste dependerá, principalmente, de la variabilidad de los textos dentro de un mismo periódico (Sánchez Carrión, 1985). Cuánto más homogéneo sea su *contenido,* menor *tamaño muestral* se precisará para representar a la *población.*

10.3. La explotación del material

A la selección de documentos *(población, muestra)* sigue la determinación de criterios para la extracción de la información en ellos "contenida". Los *criterios* varían dependiendo, fundamentalmente, de los objetivos de la investigación y de las características de los documentos. No obstante, pueden señalarse unas *pautas comunes* en los *análisis de contenido cuantitativo*. Éstas conciernen a la creación del *esquema de codificación* y la preparación del documento para su posterior tratamiento estadístico. En concreto, se diferencian tres estadios básicos en esta fase esencial al *análisis de contenido:*

a) La fragmentación del documento en *unidades de codificación* y de *contexto*.
b) La clasificación de las *unidades* en *categorías*.
c) El *registro de la información.*

10.3.1. La fragmentación del documento en unidades de codificación y de contexto

Indistintamente de cómo se vaya a analizar la información, una vez elegidos los documentos, éstos han de fragmentarse en *unidades* que ayuden a su organización y análisis.

Las *unidades* se dividen en *unidades de codificación* (o *registro)* y las *unidades de contexto*. Las *unidades de codificación* (o *registro)* se definen como "la mínima porción de contenido que el investigador aísla y separa por aparecer allí uno de los símbolos,

palabras, *slogans* o temas que el investigador considera significativos" (López Aranguren, 1989: 394). Corresponden, por tanto, "al segmento de contenido que será necesario considerar como unidad de base con miras a la categorización y al recuento frecuencial" (Bardin, 1986: 79).

Las *unidades de codificación* que se emplean con mayor frecuencia son las siguientes:

a) La *palabra*. Ésta constituye la *unidad de codificación* más básica. A ello contribuye su fácil identificación. Si bien, como indica Weber (1994), ha de tenerse presente el problema de los distintos sentidos que puede tener una misma palabra, dependiendo del *contexto* en el que se emplea. No todos los paquetes informáticos disponen de rutinas de desambiguación o de contextualización de palabras específicas.

La *codificación* de *palabras* puede incluir a todas las comprendidas en el texto, o restringirse, por el contrario, a aquéllas consideradas "claves", en conformidad con los objetivos de la investigación.

b) La *frase,* cuando se está interesado en la cuantificación de frases que contienen un mismo significado (por ejemplo, "a favor" o "en contra" de un determinado tema).

Si la *frase* es extensa o compleja, puede decidirse su división en unidades más pequeñas.

c) El *tema* en torno al cual gira el discurso del documento. Holsti (1969) lo define como la unidad textual que hace referencia a uno de los siguientes aspectos: quién percibe; el agente de la acción; la misma acción; y el objeto de la acción.

Por ejemplo, el *tema* representa una *unidad de codificación* habitual en estudios de motivaciones, opiniones, actitudes, valores, o creencias (Bardin, 1986).

d) El *párrafo*. La *codificación* de *párrafos* enteros es practicada, sobre todo, en el *análisis de contenido manual,* cuando se carece de la ayuda de programas informáticos al efecto.

Weber (1994: 265) descarta esta cuarta *unidad de codificación* por su baja fiabilidad. La experiencia de este autor le lleva a afirmar que resulta "más difícil alcanzar una fiabilidad elevada cuando se codifican unidades grandes (como párrafos), que cuando se codifican unidades más pequeñas como palabras".

e) El *texto entero,* si éste puede codificarse en su globalidad. Ello exige que el *texto* sea *breve* (titulares de periódicos, editoriales, o las respuestas a una *pregunta abierta,* por ejemplo). En caso contrario, será "difícil alcanzar una elevada fiabilidad cuando se codifican textos completos" (Weber, 1994: 265). A ello se suma otra condición importante: la idea dominante o principal ha de ser suficiente para el objetivo buscado (Bardin, 1986: 81).

f) El *personaje* o actor al que se le atribuyen las acciones que figuran en los documentos analizados.

Ésta constituye una *unidad de codificación* adecuada "siempre que el texto narre una acción" (Sánchez Carrión, 1985: 97). Bardin (1986) cita como ejemplos, las obras de ficción (films, emisiones, novelas...), que pueden analizarse siguiendo a sus personajes.

De estas *unidades de codificación,* el investigador escogerá una o varias. La elección dependerá de la conjunción de tres *aspectos* básicos:

a) Los objetivos de la investigación.
b) Las características del documento.
c) Los medios disponibles para la realización del *análisis de contenido* (manual o mediante ordenador).

Además de las *unidades de codificación,* en los documentos también se diferencian las *unidades de contexto.* Éstas son de gran utilidad para la comprensión de las *unidades de codificación.*

Bardin (1986: 82) define las *unidades de contexto* como "el segmento del mensaje cuyo tamaño (superior a la unidad de registro) es óptimo para captar la significación exacta de la unidad de registro". Si la *unidad de registro* fuera, por ejemplo, la *palabra,* la *unidad de contexto* sería la *frase;* del *tema,* el *párrafo;* del *párrafo,* el *documento entero.*

La *unidad de contexto* hace, por tanto, referencia al "contexto" donde se ubica la *unidad de codificación.* Su identificación resulta, en consecuencia, básica para la comprensión del *significado* de la *palabra,* la *frase,* o cualquiera que sea la *unidad de codificación* elegida.

10.3.2. La clasificación de las unidades en categorías

En el *análisis de contenido "categorial",* las *unidades de registro* se codifican en *categorías,* que singularicen su "contenido". Estas *categorías* proporcionan una denominación genérica a unidades de contenido similar, contribuyendo, de esta forma, a su clasificación. Para ello, las *categorías* han de cumplir los tres *requisitos* –referidos en el Capítulo 4, apartado 4.2– comunes en la medición de variables: *exhaustividad, exclusividad* y *precisión.*

No sólo han de cubrir toda la variedad del objeto de estudio. También han de evitar la "ambigüedad" en el proceso de medición, si se quiere que el *análisis de contenido* reúna condiciones de *validez* y de *fiabilidad.* Quiere esto decir que cada *unidad de registro* sólo puede clasificarse en una única *categoría.*

A esos tres *requisitos* básicos se añaden otros dos fundamentales:

a) La *pertinencia* o adecuación de las *categorías* a los *objetivos* de la investigación.
b) La *homogeneidad:* la creación de *categorías* ha de responder a un único *principio de clasificación.*

López Aranguren (1989: 398-401) diferencia los siguientes *principios de clasificación*. De ellos derivan la mayoría de las *categorías* empleadas en el análisis de material escrito:

a) *Materia, tema* o *asunto:* ¿de qué trata la comunicación?
b) *Dirección* o *punto de vista:* el tratamiento o presentación (favorable-desfavorable-neutral; positiva-negativa) que la comunicación hace de la materia o asunto del que trata.
c) *Criterio* o *valor:* sobre qué se basa la clasificación en términos de "dirección" o punto de vista. Por ejemplo, "moralidad-inmoralidad", "fuerza-debilidad".
d) *Autoridad, fuente* o *prueba:* personas, grupos, instituciones u objetos que en la comunicación se citan en apoyo de una afirmación, o en cuyo nombre se hacen declaraciones o afirmaciones.
e) *Objetivo* o *meta* que se persigue.
f) *Medio* o *método* para alcanzar los objetivos.
g) *Actores:* personas, grupos o instituciones que en el texto se representan como realizadores de determinados actos. Por ejemplo, países, grupos religiosos, económicos, etc.
h) *Características personales* y *rasgos de carácter:* edad, sexo, optimismo, depresión, entre otras.
i) *Origen:* lugares de origen de las comunicaciones.
j) *Destino,* cuando la comunicación no se dirige al público en general, sino a grupos específicos (como estudiantes, jubilados, mujeres, por ejemplo).
k) *Forma de la afirmación:* hechos, expectativas, preferencias, demandas, ...
l) *Intensidad,* cuando se clasifica el componente emocional o sentimental de las clasificaciones, el entusiasmo, la agitación o la excitación.
m) *Mecanismo* o *recurso estilístico* o *retórico,* que se emplea con fines propagandísticos, persuasivos o de tendencias.

EJEMPLO DE CLASIFICACIÓN DE UNIDADES EN CATEGORÍAS

Como ilustración de lo expuesto, a continuación se va a exponer el *sistema de categorías* desarrollado en una investigación real, en la que se aplica el *análisis de contenido*. Se trata del estudio llevado a cabo por M. Clemente y M. A. Vidal, publicado en 1994 con el título "La violencia simbólica: la televisión como medio generador de la delincuencia" (en la revista *Apuntes de Psicología,* n.° 41-42, pp. 47-60).

En esta investigación, los autores organizaron la información extraída de series de dibujos animados –emitidos en cadenas públicas (TVE1, TVE2, Telemadrid) y privadas (Tele 5, Antena 3 y Canal Plus)– en un *sistema de categorías* diseñado al efecto. Este *sistema* se confecciona conforme a tres *criterios de clasificación*:

a) Quién realiza el acto agresivo.
b) Sobre quién se comete el acto agresivo.
c) De qué acto agresivo se trata.

En los dos criterios de clasificación primeros, los autores diferencian las mismas siete *agrupaciones de categorías:*

a. Sexo 1. Varón 2. Mujer 3. Sin determinar *b. Edad* 1. Niños 2. Adolescentes 3. Adultos 4. Sin especificar *c. Apariencia* 1. Humano 2. Animal 3. Planta 4. Objeto *d. Capacidad de decisión individual* 1. La tiene 2. No la tiene	*e. Individual o grupal* 1. Uno sólo 2. Grupo 3. No está *f. Rol* 1. Protagonista 2. Habitual 3. Otros *g. Calificación del rol* 1. Bueno 2. Malo 3. Sin especificar

En cambio, en la especificación de "qué acto agresivo se realiza", los autores distinguen tres *agrupaciones de categorías:*

a. Tipos de agresiones 1. No verbales 2. Verbales leves o actos irónicos 3. Verbales graves o fuertes 4. Amenazas 5. Agresiones físicas leves (una sola) 6. Agresiones físicas fuertes (más de una) 7. Agresiones con objetos no destinados a agredir 8. Armas blancas 9. Armas	*b. Grado de proactividad del acto* 1. Proactivo 2. Reactivo *c. Grado de daño manifiesto* 1. No aparece 2. Leve 3. Fuerte 4. Muerte

La información analizada corresponde a series de dibujos animados emitidos desde el 29 de noviembre al 5 de diciembre de 1993, ambos días inclusive (tanto en días de diario, como en fines de semana). En conformidad con el *sistema de categorías* señalado, la información se cuantificó en forma de *frecuencia (absoluta* y *porcentual)* de aparición. A modo de ejemplo, en el Cuadro 10.A se recoge la tabulación efectuada para los primeros cinco *sistemas de categorías*.

CUADRO 10.A. Distribuciones de frecuencias para cinco sistemas de categorías.

		Características del agresor		*Características del agredido*	
Categoría		*Frecuencia*	*Porcentaje*	*Frecuencia*	*Porcentaje*
Sexo	1	382	64	257	43
	2	86	14	168	28
	3	131	22	174	29
Edad	1	30	4	83	14
	2	76	11	82	14
	3	271	51	230	38
	4	291	44	206	34
Apariencia	1	287	49	226	38
	2	192	33	235	39
	3	28	5	23	4
	4	79	13	111	19
Capacidad	1	461	78	387	79
decisión	2	132	22	106	21
Grupo	1	494	83	410	68
	2	93	16	190	31
	3	4	1	8	1

Una vez diseñado el *esquema de codificación,* se recomienda (véase Bardin, 1986; o Weber, 1994) su comprobación en una pequeña *muestra* (de la misma población de documentos), previa a su utilización definitiva para el registro de la información. Esta precomprobación ayudará a precisar la adecuación del *esquema de codificación* elaborado para la consecución de los objetivos de la investigación (la clasificación de las *unidades de registro,* de *contexto,* y el *sistema de categorías* creado).

10.3.3. El uso de programas informáticos para el análisis textual

Lo hasta ahora expuesto constituye el procedimiento usual, cuando el *análisis de contenido* se realiza manualmente. Sin embargo, como ya señalaba Mochmann en 1985, cada vez van adquiriendo mayor protagonismo los "procedimientos de *clasificación empírica*". En ellos, el investigador no define, *a priori,* los *conceptos teóricos.* Por el contrario, éstos resultan de la realización de procedimientos estadísticos (como el *análisis factorial* o el *análisis de conglomerados),* a partir de las *matrices de correlaciones* de palabras incluidas en los textos.

Como ilustración de este proceder, se hace mención del programa WORDS, diseñado con la finalidad de excluir cualquier influencia del investigador en la *categorización* de los textos. A tal fin, se siguen las *fases* siguientes:

a) El documento se divide en *segmentos* (páginas, párrafos, entrevistas, por ejemplo).
b) Se suprimen las *palabras funcionales* (artículos, conjunciones).
c) Se determina la *frecuencia* de aparición de cada *palabra* en cada *segmento*.
d) Se selecciona un *subconjunto* de esas *palabras* para su análisis.
e) Se calcula una *matriz de correlaciones* de las palabras incluidas en ese *subconjunto*.
f) A partir de la *matriz* se realiza un *análisis factorial*. De él se obtendrá un número de *factores*, que corresponderán a los principales *temas* de *contenido* del documento analizado (Mochmann, 1985: 37).

Asimismo, la mayoría de los programas informáticos para *análisis de contenido* no sólo proporcionan *listados de palabras*, acompañados con su *frecuencia* de aparición en el texto. También pueden proporcionar el *contexto* que arropa a la *unidad de codificación*. Esto adquiere especial relevancia en casos de ambigüedad, cuando la *palabra* presenta distintas connotaciones en *contextos* diferentes. Para su correcta clasificación en *categorías*, se crean, entonces, *índices* o *diccionarios generales*. Éstos actúan a modo de *thesaurus*, de diccionarios, que recogen los posibles sentidos de las *palabras*, en función del *contexto* en que se hallen.

El *General Inquirer* (el primero de los programas para *análisis de contenido*, desarrollado inicialmente por Stone y colaboradores, en 1966), por ejemplo, incluye varios *diccionarios generales*, especializados en áreas de conocimiento. Entre ellos se encuentran los siguientes: "Harvard Psychological Dictionary", "Stanford Political Dictionary", "Need Achievement Dictionary", "Santa Fe Anthropological Dictionary", o "Laswell Value Dictionary" (LVD).

Estos *diccionarios* proporcionan al investigador numerosas *categorías* (generalmente de 60 a 150, e incluso más), así como definiciones o reglas para la correcta clasificación de *palabras* en *categorías*. A tal fin, se considera el sentido que adopte en el *contexto* en que aparece.

Las *categorías* se representan mediante abreviaciones o *etiquetas*, que figuran en tablas. El número al final de la *palabra* indica el número de sentidos particulares de las palabras. De esta forma, la creación de *diccionarios generales* ayuda a la sistematización y estandarización del procedimiento de clasificación de las *palabras* en *categorías*.

EJEMPLO DE DICCIONARIO GENERAL

La presente ilustración se ha tomado de la obra de Bardin (1986: 100-101). Corresponde al *Stanford Political Dictionary* (de Holsti), para el análisis de documentos políticos. En él pue-

den registrarse hasta un total de 4.000 palabras, aproximadamente, según seis o cuatro conceptos clave. Por ejemplo:

Conceptos clave	*Número de palabras*	*Ejemplos de palabras*
Afecto positivo	977	Mutuo, natural, normal, puro.
Afecto negativo	1.513	Escándalo, profano, rehusar, repugnante.
Fuerza	1.391	Acero, piedra, espada, duro.
Debilidad	579	Delgado, ignorante, debajo.
Actividad	1.218	Reacción, reino, contacto, viaje.
Pasividad	722	Inmutable, espera.
Sobreestimación	128	Absolutamente, exactamente, siempre.
Subestimación	50	Aunque, aparentemente, aproximadamente
Negación	6	Diferente, ni uno ni otro.

Además de los *diccionarios,* la mayoría de los programas incluyen el procedimiento KWIC *(Key Words In Context),* que relaciona las *palabras* con el *contexto*. Por ejemplo, el programa TEXTPACK (de Mohler y Zuell, 1990, ZUMA, reléase 4.0) incluye seis operaciones básicas, entre las cuales se incluye el procedimiento KWIC y la elaboración de *índices* (o *diccionarios).* Las operaciones que incluye este paquete son:

a) *Frecuencia de palabras* diferentes para uno o más textos (procedimiento FREQ).
b) *Índice:* un listado de palabras y sus referencias en el texto (procedimiento XREFF).
c) *Palabras claves en su contexto (concordancia),* mediante el procedimiento KWIC.
d) *Comparación de vocabulario,* mediante el procedimiento WORDCOMP *(Word Comparison).* Este procedimiento está diseñado para la comparación de palabras con una denominación común o diferente.
e) *Selección de textos utilizando información externa* al texto, en forma de fichero numérico (un fichero SPSS, por ejemplo).

   En estos casos puede aplicarse el programa SUBSEL. Éste permite la selección de un texto, que corresponde a una combinación específica o valores en un fichero numérico.
f) *Codificar/categorizar palabras* que figuran en secciones concretas del texto, y su posterior inclusión en un fichero numérico (TAGCODER).

   A partir de esta información podrán después efectuarse análisis estadísticos (mediante paquetes estadísticos como el SPSS, BMDP o SAS, por ejem-

plo). Para ello se utilizarán no *palabras*, sino *grupos de palabras* clasificadas en *categorías,* en función de su significado común.

Otros programas que tan sólo se mencionan aquí son: EVA (creado para el análisis de los titulares de los diarios), ANACONDA (para la codificación de las *preguntas abiertas* de los *cuestionarios),* SPENCE (para el análisis de los *protocolos* de las *entrevistas* psiquiátricas), y el QUESTER (dirigido al análisis del *contexto)* (Mochmann, 1985).

A estos programas hay que sumar los desarrollados en fechas más recientes. Como son el LEXICLOUD (por el Laboratorio de Lexicometría y Textos Políticos de la Escuela Normal Superior de Fontenay-St.Cloud, para el establecimiento de relaciones estadísticas entre las unidades léxicas), y el sistema ALCESTE (creado por Reinert en 1990, para el análisis léxico por contexto de un conjunto de segmentos de texto). El detalle de estos programas puede encontrarse en el artículo de Julia Behar (1993).

La ayuda de estos u otros programas facilita, considerablemente, la práctica del *análisis de contenido*. Entre sus *ventajas* destacan las siguientes:

a) Pueden analizarse, simultáneamente, textos complejos, y de considerable magnitud, en un breve período de tiempo.
b) Se introduce una mayor estandarización y sistematización en el análisis, que mediante procedimientos manuales. Ello facilita la comparabilidad de los resultados analíticos.
c) Los errores de *codificación* se reducen considerablemente, al no depender ésta de la actuación del equipo investigador.

A estas *ventajas* hay que añadir la reducción en costes (económicos y temporales) que supone la aplicación de *scanners* (con reconocimiento de caracteres) para la grabación óptica de textos en el ordenador. Atrás ha quedado el tedioso procedimiento primigenio de transformación del texto en *tarjetas perforadas,* para su posterior lectura por el programa informático. Además, el uso de *scanners* presenta otra ventaja importante: la reducción de los *errores de grabación*. Los textos se introducen en el ordenador tal como aparecen en el original.

10.4. El tratamiento estadístico y la interpretación de los resultados

En el *análisis de contenido cuantitativo* se aplican las mismas *técnicas estadísticas* descritas en el Capítulo 9. A la *segmentación* de los textos sigue el *listado de frecuencias* de aparición de *palabras, frases,* o cualquiera que sea la *unidad de codificación* escogida en la investigación.

En este estadio inicial (a veces único), las *listas KWIC de palabras claves* son de gran utilidad. Estas listas muestran el *contexto* más inmediato de la *palabra,* lo que ayudará en la interpretación de su significado.

Las palabras que únicamente sirven para articular el texto (artículos, pronombres, conjunciones, preposiciones) se excluyen de los análisis. El interés, en cambio, se dirige a la cuantificación de palabras "claves"; aquellas que conforman el léxico del discurso (sustantivos, adjetivos, verbos, adverbios).

Pero, como reconoce Weber (1994: 294):

> "Aunque las listas de frecuencias de palabras revelan cambios o diferencias en el énfasis entre los documentos, deben utilizarse con precaución. Las frecuencias de palabras no revelan mucho sobre la asociación entre las palabras."

Cuando el interés del investigador va más allá de la mera descripción frecuencial de las unidades de codificación por separado, puede aplicarse cualquiera de las *técnicas multivariables* descritas en el Capítulo 9. Lo que normalmente exige la previa codificación numérica del *contenido manifiesto* del material verbal.

Una de las *técnicas multivariables* más aplicadas en el *análisis de contenido* es el *análisis factorial.* En éste, las *variables observadas* suelen ser ahora las *categorías* diferenciadas. Las *variables latentes,* por el contrario, hacen referencia a *temas* genéricos, "latentes" en el texto e, igualmente, deducidos de las *variables observadas* que conforman cada *factor* obtenido. De este modo, como en cualquier *análisis factorial,* las *categorías* que correlacionan con un mismo *factor* (es decir, aquellas que presenten un *"factor loading"* superior a .30) se interpretarán como representando un mismo *tema* en el texto analizado.

Los resultados numéricos se *validan* e *interpretan* en relación con el *tema* al que hacen referencia. Los resultados pueden corroborar las *hipótesis* de la investigación o, por el contrario, contradecirlas. Lo que puede provocar la realización de una nueva investigación.

Como ilustración de lo aquí expuesto, se han seleccionado dos investigaciones reales, en las cuales el *análisis de contenido cuantitativo* se ha materializado de forma distinta. La primera investigación ejemplifica la aplicación del *análisis factorial;* mientras que en la segunda, el análisis se restringe al recuento de *frecuencias.*

10.4.1. Ejemplos desarrollados de análisis de contenido en dos investigaciones reales

- Cabrera, J. (1992): *La nación como discurso. El caso gallego,* Madrid. CIS. *Monografía,* n.º 126.

En esta investigación se analizan distintas manifestaciones del sistema ideológico nacionalista:

a) El nivel *léxico* y *semántico,* que constituyen las estructuras superficiales del campo de la manifestación.
b) El nivel *semiológico,* que trata de desentrañar la significación profunda de dicho sistema.

Para cada nivel se ha optado por una *técnica de análisis* diferente, combinándose técnicas *cuantitativas (lexicometría)* y *cualitativas* (análisis *semántico-categorial* y análisis *semiológico).*

El *corpus* que se somete al *análisis lexicométrico* (del vocabulario que conforma el discurso) está integrado por las respuestas obtenidas mediante *entrevista abierta.* El *guión* de la *entrevista* comprende tres aspectos concretos:

a) La biografía de cada entrevistado.
b) El proceso de socialización nacionalista.
c) La posición teórico-nacionalista de los entrevistados.

La *palabra* se toma como *unidad de medida,* procediéndose a su "lematización", que comprende:

a) La *reagrupación de las formas heterogéneas de un mismo vocablo.* Consistió en reducir las formas plurales a singulares, los femeninos a masculinos, y las diversas formas de un verbo a su expresión en "infinitivo" (a excepción de los "participios". Éstos se analizaron contextualizadamente, con el propósito de determinar si constituían un adjetivo).
b) La *separación de las formas homógrafas que muestran vocablos diferentes.* Se analizaron formas con igual grafía que hacían referencia a contenidos diferentes. Estos *homógrafos* fueron identificados con un dígito, o un *dígito* y una letra, a modo de *sufijos.*

   Por *ejemplo,* la forma adjetiva de "gallego" se distinguió de sus formas sustantivas, quedándose como sigue:

   - "Gallego" = adjetivo
   - "Gallego2" = sustantivo
   - "Gallego2a" = sustantivo referido al idioma

   En general, se adoptaron las claves siguientes: "2" para sustantivos; "2a" para sustantivos referidos a idiomas; "3" para adjetivos; "4" para adverbios; y, "5" para aquellos nombres propios de personas o ciudades que podían inducir a confusión.

Finalizada la *lematización,* se efectúa el recuento mecánico de los textos. Para ello se diferencia, a nivel *léxico* (de vocabulario), dos tipos diferentes de *palabras:*

a) *Palabras de relación o gramaticales,* aquellas que por sí mismas carecen de significado: artículos, pronombres, preposiciones y conjunciones.
b) *Palabras lexicales* o formas plenas de significado. Éstas conforman el *léxico* del discurso: sustantivos, adjetivos, adverbios y verbos.

La investigación se centra en el "discurso lexical", en el análisis estadístico de los rasgos de contenido. En el Cuadro 10.2 figura la distribución general de frecuencias para cada uno de los 12 sujetos cuyos discursos forman la *muestra* del estudio. Los diversos "subcorpus" se distribuyen en función de la extensión del:

a) *Número de ocurrencias* (N): el número de ítems. Refleja, por tanto, su extensión.
b) *Vocabulario* (V): número de formas diferentes empleadas.
c) *Hapax:* aquellas formas cuya frecuencia de aparición es igual a 1.
d) *Número de ocurrencias funcionales* ($FF_{(n)}$): items que se corresponden con el vocabulario funcional ($FF_{(v)}$).
e) *Número de ocurrencias lexicales* ($FL_{(n)}$), con la extensión del vocabulario correspondiente.

Cada uno de estos valores figura acompañado de su correspondiente *valor de rango* (r) para cada columna. El *corpus* total (suma de los diversos *subcorpus* correspondientes a cada emisor) tiene una extensión de 133.249 palabras. Éstas se distribuyeron porcentualmente en los diversos *subcorpus* como se refleja en el Cuadro 10.1.

CUADRO 10.1. Distribución general de frecuencias.

	N	r	V	r	Hapax	r	$FF_{(n)}$	r	$FF_{(v)}$	r	$FL_{(n)}$	r	$FL_{(v)}$	r
GM	5524	1	1113	1	613	2	3066	1	162	2	2458	1	951	1
VA	6521	2	1151	2	598	1	3766	2	160	1	2755	2	991	2
FE	8378	3	1601	7	875	8	4624	3	175	5	3754	3	1426	7
GS	9593	4	1479	3	762	4	5543	5	183	8	4050	5	1296	3
JB	9605	5	1580	5	821	7	5314	4	177	6	4291	6	1403	5
CH	9763	6	1503	4	771	5	5722	6	171	3	4041	4	1332	4
PA	11802	7	1581	6	698	3	6618	8	174	4	5184	7	1407	6
CC	12098	8	1869	10	959	10	6556	7	189	9	5542	10	1680	10
LS	12655	9	1751	9	924	9	7218	9	190	10	5437	8	1561	9
RA	13359	10	1695	8	792	6	7845	10	178	7	5514	9	1481	8
LF	16794	11	2206	11	1136	11	9166	11	198	12	7628	11	2008	11
BA	17157	12	2284	12	1158	12	9492	12	192	11	7665	12	2092	12
Total	133249				1010		74930 %N: 56,23				58319 %N 43,77			

En el Cuadro 10.2 destacan los discursos de los sujetos BA y LF, cuya extensión supera a la del resto de los entrevistados. Ambos acumulan algo más de la cuarta parte de todo el *corpus*. Los discursos de RA, LS, CC y PA forman un segundo bloque. Sus discursos superan el millar de *ítems,* pero no alcanzan a superar el millar quinientos mil. Por lo que, acumulan el 37,5% del total.

Respecto a la *lexicalidad,* en esta investigación se confirma una norma observada en la generalidad de los *análisis de contenido cuantitativo:* las *formas funcionales* (FF) superan el 50% de la extensión general. En este caso, el 56%. En la investigación se analizará, en consecuencia, el 44% restante que conforma el "discurso lexical": las formas que caracterizan el discurso sobre la nación.

A *nivel léxico* se diferencia entre:

a) *Vocabulario original:* toda forma que sólo se encuentra en un emisor de la *muestra.*
b) *Vocabulario común:* formas que tienen al menos una ocurrencia en todos los emisores.
c) *Vocabulario de base:* formas que tienen una distribución regular en todos los emisores. Por lo que no introducen diferencias.
d) *Vocabulario específico:* el que caracteriza a cada uno de los emisores.

El subconjunto de formas así obtenido constituye el *corpus* sobre el que se realizan los análisis. El propósito es determinar las especificidades de cada emisor.

Para el cumplimiento de dicho objetivo, se aplica un *análisis factorial de componentes principales* mediante el paquete estadístico BMDP. La *matriz de correlaciones* la forman cada uno de los 12 discursos que componen la *muestra* del estudio. Estos *discursos* actúan como *variables* y las 347 *palabras* como *casos.* De lo que se trata es de analizar las interrelaciones existentes entre ellos, en busca de semejanzas y especificidades.

La *matriz de correlaciones* entre las 12 variables se muestra en el Cuadro 10.2.

CUADRO 10.2. Matriz de correlaciones.

	1	2	3	4	5	6	7	8	9	10	11	12
1	1.00											
2	–.04	1.00										
3	–.10	–.06	1.00									
4	–.04	–.07	–.04	1.00								
5	–.12	–.05	–.03	–.03	1.00							
6	–.00	.05	–.07	–.08	–.10	1.00						
7	–.05	–.11	–.06	–.13	–.07	–.05	1.00					
8	–.05	–.10	–.09	–.16	–.12	–.05	.11	1.00				
9	–.14	.00	–.00	–.08	–.13	–.06	–.13	–.11	1.00			
10	–.15	–.10	–.03	.07	.06	–.19	–.25	–.30	.05	1.00		
11	–.04	–.09	–.15	–.10	–.10	–.13	–.20	–.20	–.12.	–.08	1.00	
12	–.12	–.06	–.08	–.17	–.10	–.12	–.17	–.25	–.13.	–.06	–.08	1.00

La *matriz factorial* de los 4 *factores* que resultan de la agrupación de las 12 variables, en función de la semejanza de los discursos, se rota por el procedimiento más habitual: *rotación ortogonal varimax*. Se busca la disposición de los *componentes* o *factores* de manera que cada uno de ellos tuviera saturaciones elevadas en pocas variables (en este caso, emisores de los discursos), y muy bajas en el resto.

El Cuadro 10.3 muestra las *cargas factoriales* de cada variable en los 4 *factores*. Todas las variables presentan, al menos, un *"factor loading"* superior a .25, en alguno de los *factores*. Como es usual en este tipo de análisis, los *factores* extraídos aportan de más a menos información: el primer *factor* explica más *varianza* (14%) que el segundo (11%), y éste más que el tercero (10%). La *varianza* explicada por el tercer *factor* es ligeramente superior a la explicada por el cuarto. Véase el Cuadro 10.3.

CUADRO 10.3. Matriz factorial rotada oblicuamente mediante el procedimiento varimax (*factor loadings* superiores a .25).

Emisor	Factor 1	Factor 2	Factor 3	Factor 4
GM	–	.517	–	–.394
VA	–	–	.570	–
FE	–	–.432	–	–
GS	–.268	–	–	–.681
JB	–	–	–.476	–
CH	–		.629	–
PA	.690	–	–	–
CC	.724	–	–	–
LS	–	–.562	–	–
RA	–.563	–.271	–.395	–
LF	–.253	.604	–	–
BA	–	–	–	–.726
Varianza	13,78%	10,82%	10,33%	10,06%

Para cada *factor* se presentan las formas (o *palabras)* positivas y negativas que lo componen, y sus emisores. En el Cuadro 10.4 se extractan algunas de las *palabras* correspondientes al *factor* 1. Este *factor* está caracterizado por los emisores CC, PA y RA, con unos *pesos* de .724, .690 y .563, respectivamente.

En los *valores positivos* se encuentran la mayoría de las *formas* que caracterizan tanto a CC como a PA. Destacan, especialmente, las correspondientes a los *ámbitos lingüísticos* ("lingua", "catalán2a", "catalán", "portugués2a", "ortografía", etc.), con resonancias de elementos discursivos del galleguismo de la época republicana.

El lado *negativo* del primer *factor* está fuertemente dominado por RA. En él se encuentra un léxico de corte más político, de claras referencias partidistas y electo-

rales ("PSOE", "CiU", "voto", "OTAN", "votar", "AP"); junto a un tímido eco del léxico sociológico y de estratificación social ("social", "clase", "sector", "sociedades"). En estos ámbitos el entrevistado RA mostraba una clara especialización.

CUADRO 10.4. Extracto de algunas de las formas positivas y negativas que componen el factor 1.

Factor 1				Emisores
		Formas positivas		
lingua	3.490	hispanico	1.540	CC .724
catalán2a	3.100	protagonismo	1.540	PA .690
catalán	2.750	mesmo	1.530	
portugués2a	2.710	galeguista	1.490	
ortografía	2.620	universal	1.470	
mentalidade	2.580	forza	1.460	
		Formas negativas		
pensar	–3.928	problema	–1.200	RA –.563
cousa	–2.078	caer	–1.183	
socia	–1.900	formulación	–1.122	
PSOE	–1.397	votar	–1.027	
diferente	–1.315	AP	–1.000	
OTAN	–1.209	traballar	–1.000	

A partir de las *formas positivas* y *negativas* que configuran cada *factor,* se da a cada uno de ellos las "etiquetas" o denominaciones siguientes:

- Factor 1: cultural-liberal / político-partidista
- Factor 2: sociológico-administración / político– estratificación social
- Factor 3: cultural-identitario / histórico-social
- Factor 4: estratégico-economía política / dispersión político-cultural

Los 4 *factores* se cruzan entre sí (el 1.° con el 2.°; el 1.° y el 3.°; el 1.° y el 4.°, y así), con la finalidad de comprobar la diferenciación léxica entre las distintas generaciones de nacionalistas entrevistados. Cada uno de los cruces se representan gráficamente mediante *planos de proyecciones* de las *palabras* que forman cada *factor*. De esta manera se obtienen las *estructuras léxicas* de las manifestaciones ideológicas.

El primer estadio de la investigación (el correspondiente al *análisis de contenido cuantitativo)* se da por concluido. Para un mayor conocimiento de esta primera etapa de investigación, así como el desarrollo del *análisis cualitativo,* remito a la publicación de esta excelente investigación de Julio Cabrera.

- Llovet, J. J. (1992): "El control de la prensa sobre la profesión médica: el caso de EL PAÍS", *Revista Española de Investigaciones Sociológicas,* n.º 59, pp. 261-285.

En esta segunda investigación se estudia el control social sobre la medicina a través de la prensa. Para ello se realiza un *análisis de contenido* centrado en el diario *El País*.

La *muestra* abarca las noticias, cartas de lectores, editoriales, etc., publicados en el lustro 1985-1989, referidos a negligencia y conductas médicas incorrectas o lesivas. Se descartaron, por el contrario, las informaciones que aludían a hechos o procesos de este carácter ocurridos fuera de España; al igual que los derivados de la atención psiquiátrica (salud mental), por ser ésta un área muy específica.

En primer lugar se lleva a cabo un rápido peinado visual de los titulares (antetítulos y títulos), aparecidos en las secciones del diario en las que corrientemente pueden leerse comentarios o información sobre el tema. Pocos meses del año quedan en blanco, sin esta clase de información.

En el Cuadro 10.5 se exponen las *distribuciones porcentuales* de las unidades de redacción de *El País*. Éstas se refieren a negligencias y actuaciones incorrectas, o lesivas, de médicos, según la subsección del diario y género periodístico (del 1 de enero de 1985 al 31 de diciembre de 1989).

CUADRO 10.5. Distribuciones porcentuales de unidades de redacción de *El País*.

Frecuencia porcentual		
Subsección del diario		
"Opinión"/"Cartas al director"		22,9
"Sociedad"/"Tribunales"		22,2
"Sociedad"/"Sanidad"		21,5
"Sociedad" (sin especificación)		15,6
"España"		7,3
"Madrid"		5,2
"Opinión"/"Editorial"		2,8
"Domingo"		1,0
"Sociedad"/"Ciencia"		0,3
"Contra-tapa"		0,3
"Opinión"/"El *Ombudsman*"		0,3
"Opinión"/"Fe de errores"		0,3
	Total	100 (288)
Género periodístico		
Noticia		70,1
Carta de lector		22,9
Editorial		2,8
Reportaje, informe, reseña		2,8
Artículo o columna firmada		1.0
Errata		0.3
	Total	100 (288)

También se calculan las *frecuencias porcentuales* de las noticias, según el papel social o institucional de los enunciadores: jueces o magistrados, médicos denunciados, familiares o amigos de los pacientes perjudicados, pacientes perjudicados, abogados de los pacientes, colegios de médicos, etc.

De toda la información extraída del periódico *El País,* se deduce la contribución de la prensa como factor preponderante en la apertura, esclarecimiento y sensibilización de los ciudadanos, y pacientes, hacia la problemática sanitaria.

Como puede verse, el uso que del *análisis de contenido* se ha hecho en esta segunda investigación es bastante más simple que el efectuado en la primera investigación. Se limita a la descripción de frecuencias y comentario de extractos de textos que aparecieron en prensa durante el período analizado. No obstante, el principal interés en la exposición de esta segunda investigación es ilustrar un modo de hacer *análisis de contenido,* también usual en la investigación social.

10.5. La aplicabilidad del análisis de contenido y los límites a la inferencia

La aplicación del *análisis de contenido* se ha vinculado, tradicionalmente, al estudio de la comunicación. En 1969 Holsti sintetiza su *aplicabilidad* en la resolución de los siguientes interrogantes de la comunicación:

a) Qué se dice:

- Describir tendencias y diferencias en el "contenido" de la comunicación.
- Relacionar las características de las fuentes con los mensajes que éstas generan.
- Evaluación del "contenido".
- Extraer los patrones culturales presentes en el mensaje.

b) Cómo se dice:

- Comprobar si en el mensaje se hace uso de técnicas persuasivas.
- Analizar el estilo del texto.

c) A quién:

- Qué peculiaridades presentan aquéllos a quienes está destinado el mensaje.

d) Por qué:

- Identificar las intenciones y otras características de los emisores del mensaje.

e) Quién lo dice:

- Determinar la autoría del emisor.

f) Con qué efectos:

- Analizar el flujo de la comunicación.

Sin embargo, la aplicación que del *análisis de contenido* se hace desde la vertiente metodológica cuantitativa se halla condicionada por *limitaciones* importantes:

a) Aunque la *cuantificación* ayude a la sistematización y estandarización del análisis textual, no siempre se profundiza en los *contenidos no manifiestos* de los mensajes.
b) La interpretación de las *unidades de codificación* no ha de restringirse al mero recuento de *frecuencias* de aparición. Por el contrario, exige su previa contextualización. Precisamente, porque unas mismas palabras, en contextos diferentes, pueden implicar significados distintos.

Estas limitaciones pueden solventarse complementando el *análisis de contenido "categorial"* con el *"contextual"*, el *"cuantitativo"* con el *"cualitativo"* (este último más centrado en la indagación de las *estructuras semánticas* y *semiológicas* de los documentos). Como bien reconocen Ruiz Olabuenaga e Ispizua (1989: 185), "los mejores análisis de contenido son aquellos que utilizan la técnica de la *triangulación*". Ello permite una mejor captación de su pleno significado, la definición de la situación y del punto de vista del emisor.

10.5.1. Cuestiones de validez y de fiabilidad

La *inferencia* en el *análisis de contenido* está condicionada, como en cualquier procedimiento analítico, a la *validez* y *fiabilidad* de la *medición* aplicada. En este caso, la síntesis hecha de los "contenidos" de los documentos que se analizan.

La *validez* indica la *adecuación* del sistema de clasificación seguido en la medición de los *conceptos* teóricos; mientras que la *fiabilidad* expresa la *consistencia* del procedimiento de medición.

• Respecto a la *fiabilidad,* han de arbitrarse unas reglas de *codificación* que permitan que distintos codificadores alcancen, independientemente unos de otros, los mismos resultados. Esto es más fácilmente alcanzable cuando el *análisis de contenido* se realiza mediante ordenador. Una vez que las *reglas de codificación* se han definido, éstas se aplican automáticamente, y de igual manera en cada uno de los casos analizados.

Krippendorff (1980, 1990) diferencia tres tipos de *fiabilidad* pertinentes en el *análisis de contenido:*

a) *Estabilidad:* grado en el que varían los resultados de la clasificación a lo largo del tiempo.

Deberían alcanzarse los mismos resultados en la clasificación de las *unidades de codificación,* siempre que se emplee la misma agrupación de *categorías*. Ello exige la eliminación de cualquier ambigüedad en la clasificación de las *unidades* en categorías.

b) *Reproducibilidad:* grado al que se alcanzan los mismos resultados, cuando un mismo texto se codifica por varias personas.

En este caso, la *infiabilidad* se atribuye, sobre todo, a arbitrariedades en la *codificación*. Estas pueden deberse a ambigüedades en las instrucciones de codificación. Pero, también a errores de grabación o a diferencias cognitivas entre las personas que realizan la *codificación*.

c) *Adecuación:* grado en el que la clasificación del documento se corresponde con un estándar o norma.

Como rara vez se establecen *codificaciones estándar,* este último criterio de *fiabilidad* apenas se aplica.

Para la cuantificación de la *fiabilidad* se han propuesto algunos *coeficientes*. Sánchez Carrión (1985) destaca los dos siguientes:

a) La *razón de codificaciones de acuerdo,* entre el total de codificaciones:

$$C.R. = \frac{2M}{N_1 + N_2}$$

donde: "M" corresponde al número de codificaciones en las que coinciden los 2 codificadores.
"N_1" y "N_2", el número de codificaciones efectuadas.

b) El *índice de fiabilidad* (p_i) de Scott:

$$p_1 = \frac{\text{proporción de acuerdo observado} - \text{proporción de acuerdo esperado}}{1 - \text{proporción de acuerdo esperado}}$$

Este índice oscila entre .00 y 1.00. El valor 1.00 representa el acuerdo total entre los codificadores; mientras que .00, el desacuerdo.

La *validez,* en cambio, denota correspondencia entre los *constructos teóricos* y las *variables empíricas* empleadas. Esto exige, igualmente, la eliminación de cualquier ambigüedad en la clasificación de las *unidades de codificación* en *categorías*.

En concreto, se diferencian varios tipos de *validez* en el *análisis de contenido* (López Aranguren, 1989; Krippendorff, 1990; Weber, 1994):

a) *Validez pragmática* o de utilidad de los resultados del análisis en el estudio del tema que se investiga.

b) *Validez de constructo:* correspondencia entre la medida aplicada con alguna otra utilizada del mismo *constructo*.

c) *Validez de contenido:* grado en el que los *indicadores* cubren la variedad de significados incluidos en el *concepto*.

d) *Validez predictiva:* grado de acuerdo/correspondencia entre las predicciones obtenidas y los hechos observados.

Las *inferencias* hechas a partir del análisis de la información han de "mostrar un alto grado de acuerdo con los estados, atributos, sucesos o propiedades del contexto de los datos al que esas inferencias se refieren, y un alto grado de desacuerdo con las características contextuales que esas diferencias excluyen lógicamente" (Krippendorff, 1990: 232).

e) *Validez semántica:* grado en el que las *unidades de codificación* clasificadas juntas poseen connotaciones similares.

En el *análisis de contenido* "se logra una alta validez semántica cuando la semántica del lenguaje de los datos se corresponde con la fuente, el receptor o cualquier otro contexto respecto al que se examinan dichos datos" (Krippendorff, 1990: 231).

Los límites a la *inferencia* se hallan, en conclusión, en la ambigüedad de los significados de las *palabras* y de las *reglas de codificación*. Por lo que se recomienda su explícita delimitación.

Lecturas complementarias

Bardin, L. (1986): *Análisis de contenido*. Madrid, Akal.

Behar, J. (1993): "Aproximación al análisis textual informatizado". *Anuario de Psicología,* n.º 59, pp. 61-78.

Clemente Díaz, M. (1992): "El análisis de contenido". En Clemente, M. (comp.): *Psicología Social: métodos y técnicas de investigación*. Madrid, Eudema, pp. 169-207.

Krippendorff, K. (1990): *Metodología del análisis de contenido: teoría y práctica*. Barcelona, Paidós.

López-Aranguren, E. (1994): "El análisis de contenido". En García Ferrando, M. *et al.* (comp.): *El Análisis de la Realidad Social*. Madrid, Alianza, pp. 461-492.

Mochmann, E. (1985): "Análisis de contenido mediante ordenador aplicado a las ciencias sociales". *Revista Internacional de Sociología,* vol. 43 (1), pp. 11-44.

Sánchez Carrión, J. J. (1985): "Técnicas de análisis de los textos mediante codificación manual". *Revista Internacional de Sociología,* vol. 43 (1), pp. 89-118.

Weber, R. P. (1994): "Basic content analysis". En Lewis-Beck, M. S. (comp.) *Research Practice*. California, Sage, pp. 251-338.

EJERCICIOS PROPUESTOS

1. Describa una investigación en la que se haya hecho un análisis de contenido cuantitativo.
2. Analice el contenido de los editoriales del próximo fin de semana que aparezcan en tres periódicos.
3. ¿Qué peculiaridades presentan los criterios de validez y de fiabilidad en el análisis de contenido? Señale semejanzas y diferencias con los enunciados en el Capítulo 4.
4. Un grupo de investigadores desea analizar la influencia de la prensa escrita en el desarrollo de una campaña electoral. Para ello deciden analizar los contenidos de periódicos publicados durante la última campaña electoral. Diseñe la investigación: especifique cómo se seleccionaría la muestra del estudio, cuáles serían las unidades de codificación y de contexto, y trace el plan de análisis.

11

EL INFORME DE LA INVESTIGACIÓN

La investigación sociológica culmina con la redacción del *informe*. En él cristalizan cada uno de los estadios que forman el proceso investigador: desde la formulación inicial del problema y la delimitación de los objetivos, hasta la interpretación de los resultados de la investigación. Sin esta exposición pormenorizada del estudio (de su desarrollo y del alcance de los hallazgos) no podría darse por concluida la investigación. El investigador estaría más lejos de poder cumplir uno de los requisitos fundamentales de su trabajo: la comunicación pública en general y a la comunidad científica en particular, para su conocimiento y valoración. Como Caplow (1977: 15) afirma:

> "El objetivo oficial del investigador es siempre el de contribuir al enriquecimiento del conocimiento científico. Se sobrentiende que aquél publicará los resultados obtenidos sin restricciones ni deformaciones, y que proporcionará suficientes informaciones sobre sus métodos a fin de que otros investigadores puedan reproducir este estudio si así lo desean."

El *informe* adquiere tal relevancia en el esfuerzo investigador, que Caplow (1977) llega incluso a negar la consideración de "investigación sociológica" a aquélla cuyos resultados no sean publicados.

A continuación se exponen algunas consideraciones genéricas, a tomar como referencia, en la elaboración de un *informe* de investigación.

11.1. La organización del informe: contenidos básicos

El *informe* cumple la función de presentación, de comunicación, de un esfuerzo investigador. Razón por la cual ha de referirse explícitamente a las distintas fases que

componen la investigación, desde su concepción inicial hasta su conclusión. Se recomienda que esta exposición sea objetiva, transparente, y lo más fidedigna posible respecto a los contenidos de la investigación. En palabras de Alvira (1995: 246-247):

> "El informe tiene que dar cuenta del cómo, cuándo y con qué resultados se han cumplido los objetivos del proyecto. Un informe debe explicar los procesos seguidos, las divergencias o cambios en lo proyectado (si los hubiere) y la confirmación o refutación de las hipótesis de partida que dieron lugar al proyecto de investigación."

Representa, en suma, la *síntesis,* la culminación del ejercicio investigador, en forma de documento para su lectura y evaluación. Este documento puede adoptar distintas formas en conformidad con las características del público al que principalmente se dirige (profano, experto), del medio empleado en su divulgación (monografía, revista, comunicación de congreso, informe técnico), y del ingenio y estilo comunicativo del investigador, fundamentalmente.

Dependiendo de cómo se describa la investigación (genérica o pormenorizadamente), de qué aspectos se antepongan (los metodológico-técnicos al comentario de resultados), y del grado de tecnicismo que se dé a la redacción de los contenidos de la investigación, se diferencian varios tipos de informes. Desde los *técnicos* (dirigidos a una audiencia interesada en el conocimiento de todo el desarrollo metodológico de la investigación) a los *divulgativos* (centrados en el comentario de los resultados de la investigación, sin tecnicismos, en un lenguaje accesible al público general).

No obstante, en la mayoría de los informes de investigación puede encontrarse una *estructura básica común,* que adquirirá distinto desarrollo en función de las características del *informe.* Esta *estructura* está integrada por seis grandes apartados:

a) Introducción.
b) Metodología.
c) Resultados.
d) Discusión.
e) Conclusión.
f) Bibliografía.

• *Introducción*

El *informe* suele comenzar con una sección introductoria, que ofrece una visión global de la investigación y su relevancia para el conocimiento del problema que se analiza. El investigador define y contextualiza el *problema* de investigación, especificando los *objetivos* e *hipótesis* esenciales en la indagación. Pero también resalta la trascendencia del esfuerzo investigador realizado respecto a indagaciones anteriores.

En su justificación de la investigación, el investigador destaca el interés de su estudio y por qué constituye éste un avance en el conocimiento del problema que ana-

liza. Para ello contextualiza la investigación en su área de conocimiento. El propósito no es mostrar un catálogo de la literatura especializada en el tema, sino señalar cuál es el estado actual de la cuestión: qué cuestiones se han resuelto y sobre cuáles existen aún controversias o una falta de estudios.

Dicha argumentación se acompaña de un número limitado de referencias (de estudios de notorio interés), seleccionadas para marcar la trayectoria investigadora precedente.

• *Metodología*

Bien después de la "introducción", bien al final del *informe,* en un "apéndice metodológico", se describen, con mayor o menor detalle (dependiendo del *informe),* los aspectos relacionados con el *diseño de la investigación.* Concretamente, se expone qué *estrategia* se siguió en la indagación del problema de estudio; se datallan las características técnicas de la *muestra;* y se ofrecen comentarios, que se juzguen pertinentes, sobre las *técnicas de obtención* y de *análisis* de datos empleadas. Además, se comentan las incidencias (si las hubiera) en la materialización del *diseño* proyectado, señalándose las actuaciones para su resolución.

El propósito fundamental de este apartado –como indica Gilbert (1993: 338)– reside en "proporcionar el detalle suficiente para que otro investigador pueda repetir la investigación". Por esta razón han de explicitarse cada uno de los pasos seguidos en la ejecución del *proyecto* de investigación: la operacionalización de los *conceptos teóricos* fundamentales, la selección de la *muestra,* la fecha del *trabajo de campo,* el porcentaje de respuestas, junto con la descripción de las *técnicas* utilizadas para la *recogida* (cuestionario, guión de entrevista...) y *análisis* de la información.

Además de describir las singularidades del *diseño* de la investigación, es importante que se indiquen las posibles limitaciones de la información obtenida: su *validez* y *fiabilidad.*

Por *ejemplo,* Fowler (1988: 143) indica, en el caso de una *encuesta,* la siguiente información a incluir en el *informe* de la investigación:

a) El *marco de muestreo* y la estimación del porcentaje de la *población* que tuvo oportunidades de selección. A ser posible, también señalar si la *población* excluida de la *muestra* difiere de aquella que finalmente participa en el estudio.
b) El *procedimiento de muestreo* seguido en la selección muestral.
c) *Resultados del campo:* porcentaje de respuestas y razones principales para la "no respuesta".
d) Descripción de *quién llevó a cabo la entrevista:* si se contrataron entrevistadores y supervisión efectuada.
e) Breve descripción del *diseño del cuestionario,* incluyendo su precomprobación *(pretest).*
f) La *redacción exacta de las preguntas,* si éstas no están reproducidas en el texto.

El *informe* debería incluir el *cuestionario* íntegro, al menos en un anexo al final del documento. El conocimiento de la redacción exacta de las preguntas y orden de las mismas, sin duda, ayudará al lector en la valoración de la información publicada en el *informe.*

El *cuestionario* suele figurar al final del documento, acompañado de las frecuencias obtenidas en cada opción de respuesta de las preguntas.

g) La calidad de los procedimientos de comprobación y control durante la *codificación, grabación* y *análisis* de los datos.

h) Un breve comentario sobre los posibles *errores* en la *encuesta.*

• *Resultados*

Descrita la metodología, se pasará en el *informe* a la relación y comentario de los "hallazgos" de la investigación. La atención se pone, sobre todo, en los *resultados* más relevantes respecto a los *objetivos* del estudio.

La exposición de los *resultados* ha de caracterizarse por su claridad, precisión y objetividad. El texto puede acompañarse con *gráficos* y *tablas* que faciliten la exposición y la lectura de los resultados numéricos del estudio.

• *Discusión*

Los *resultados* se interpretan relacionándolos con las *hipótesis* iniciales de la investigación. Se discute si éstas son corroboradas por la información reunida.

En esta *discusión* también se comparan los *resultados* de la investigación con los de estudios (si los hubiese) anteriores, además de precisar la relevancia y los límites a su generalización.

De toda esta reflexión se deducirá la respuesta de algunas de las cuestiones planteadas en la *introducción* del *informe.* Otras, sin embargo, quedarán sin respuesta, poniendo de manifiesto las limitaciones del estudio. Por lo que es común que este apartado concluya con la propuesta de futuras líneas de investigación.

• *Conclusión*

De todas los apartado del *informe,* éste es –de acuerdo con Gilbert (1993: 340)– el "más difícil de escribir porque debería resumir cuáles fueron los principales hallazgos de la investigación de una forma sucinta e interesante. Esto es, porque muchos lectores mirarán primero el final para ver si tiene algo interesante que decir, antes de comenzar por el principio".

Se relatan los hallazgos más significativos del estudio, dejando la puerta abierta a futuras indagaciones que resuelvan cuestiones discutidas o sin respuesta en la investigación descrita.

• *Bibliografía*

A lo largo del *informe* se habrá hecho referencia a otros estudios de interés, sobre el mismo *problema* o problemas colaterales. Estas referencias figuran en el texto como *citas*, con el primer apellido del autor, seguido de la fecha de publicación de la obra, todo escrito entre paréntesis, generalmente. Si de la obra se ha extraído literalmente un fragmento, a la *cita* se adjunta el número de las páginas en que dicho fragmento se halla. Esto permite que el lector pueda localizar la *cita* en su contexto original. Pero, para ello debe previamente conocer la referencia completa de la obra.

Al final del *informe* figuran las *referencias bibliográficas* completas de los textos consultados en la investigación, normalmente ordenadas alfabéticamente por autor.

En las *referencias* de *libros* se escribe el apellido principal del autor y las iniciales del nombre, la fecha de publicación de la obra (entre paréntesis), el título del libro (generalmente en cursiva), la ciudad donde se publica y el nombre de la editorial. Para las *revistas,* el título del artículo aparece entre comillas, acompañado del nombre de la revista (en cursiva), el número o volumen de la misma, y las páginas correspondientes al artículo.

En la *bibliografía* han de figurar, únicamente, las referencias exactas de aquellos estudios a los cuales se ha hecho mención en el *informe*. Por tanto, se trata de incluir sólo las obras realmente consultadas en la realización y redacción del estudio.

A este esquema genérico de *informe* a veces se añade un breve *resumen,* que contiene la información más notoria de la investigación. Este *resumen* puede figurar al principio o al final del *informe*.

Si el *informe* presenta la forma de *artículo* de revista, por ejemplo, el *resumen* aparece al principio del artículo: después del *título* y antes de la *introducción*. Su extensión es muy reducida. En un máximo de 10 líneas, suele resumirse el contenido principal del artículo. Este breve *resumen* cumple la función de ofrecer una visión sintética y global de la investigación, con el propósito de captar la atención del lector y despertar su interés por su lectura.

También es frecuente dedicar unas líneas, al comienzo del *informe,* a los *agradecimientos* o mención de las personas e instituciones que contribuyeron en la realización de la investigación: entidades que la financiaron o colaboraron de algún modo en la materialización del estudio; y particulares que ayudaron al equipo investigador durante el curso de la investigación o en su conclusión (en la lectura del borrador del *informe*, por ejemplo).

La *estructura del informe,* aquí expuesta, puede experimentar variaciones según la especificación y complejidad que el investigador dé a su *informe*. Sobre todo dependerá de la formación e interés del lector al que el *informe* se dirige principalmente.

> "Un informe escrito como un documento privado para el departamento o empresa que financia la investigación enfatizará cosas diferentes, quizás dando menos espacio a la metodología, que uno dirigido a expertos. Ambos serán más formales y técnicos que un informe dirigido al público general" (Moser y Kalton, 1989: 467).

11.2. La redacción del informe

La redacción del *informe* supone el esfuerzo de plasmar y sintetizar buena parte de las tareas llevadas a cabo durante la investigación. La escritura de lo hecho y su justificación puede resultar para algunos sencilla, mientras que para otros se convierte en una labor tediosa. Dependerá de la práctica del investigador, de su capacidad creativa, así como de su habilidad en las artes de la comunicación escrita.

En este último apartado lo único que se pretende es señalar algunas sugerencias que ayuden a afrontar la redacción de un *informe* de investigación.

Para la *redacción del informe* es muy importante que el investigador identifique, previamente, las características e intereses de su *audiencia.* Los rasgos de ésta habrán de tenerse en cuenta a la hora de dar forma al *informe.* Reléase la cita de Moser y Kalton, reproducida más arriba.

En concreto, de la *audiencia* interesa conocer:

a) Su grado de conocimiento sobre el tema expuesto.
b) El lenguaje en que se expresa.
c) Las cuestiones que más le interesan.

Este conjunto de aspectos llevará al investigador a escribir de una forma más técnica o, por el contrario, más coloquial; a exponer los contenidos de la investigación de una forma más detallada o más genérica; a poner mayor énfasis en la metodología o en los resultados.

> "Si se escribe para científicos sociales, pueden hacerse determinados supuestos sobre su conocimiento previo, y quizás puedan resumirse ciertos puntos más que explicarse en detalle. Igualmente puede emplearse apropiadamente un lenguaje más técnico que el apropiado para una audiencia general. Términos y supuestos aceptables para los colegas próximos pueden sólo confundir a otros científicos" (Babbie, 1992: A9).

Si bien, siempre se ha de procurar facilitar la lectura del *informe.* De ahí la usual recomendación de traducir los *tecnicismos* a un lenguaje conocido por el lector potencial de la investigación. La *redacción* ha de ser lo más clara y accesible posible.

> "No importa para quién estás escribiendo, mantén tus frases cortas y sencillas, con preferencia a los términos técnicos. Las construcciones complejas sólo confunden al lector (y podrían llevar a la sospecha de que tú, el escritor, no estás pensando con claridad tampoco)" (Gilbert, 1993: 334).

Para la *redacción del informe* también es importante la identificación del *medio* que va a utilizarse para la difusión de la investigación. Si va a publicarse dentro de una colección de monografías o en una revista científica, resulta conveniente la consulta previa de algunos de los números editados con anterioridad. De su lectura, el investiga-

dor podrá conocer qué formato se espera que cumpla su *informe:* su estructura, extensión, y redacción de los contenidos.

Identificada la *audiencia* y el *medio,* el investigador puede comenzar la *redacción del informe.* Para ello tan importante o más que lo anterior es la previa elaboración de un *esquema* o *índice,* que reúna los principales puntos a cubrir en el *informe.*

Este *esquema* suele ajustarse a la estructura general expuesta en el apartado anterior, en correspondencia con las distintas fases del *proyecto de investigación.* Con él se pretende organizar la información de manera que se facilite su escritura.

El investigador puede comenzar por el primer apartado y seguir el orden del *índice* hasta el final. Pero, puede también actuar siguiendo otros órdenes y cubrir a saltos el *esquema* inicial.

Una práctica usual es dejar la redacción de la *introducción* y la *conclusión* para el final. Como ambos apartados ofrecen una visión conjunta de la investigación, lo más práctico e inteligente es aplazar su redacción para el final. Si se opta por lo opuesto, la redacción de la *introducción* puede verse alterada tras la escritura del *informe* completo.

También, conviene tener presente que la *introducción* y la *conclusión* suelen ubicarse al principio y al final del *informe,* respectivamente. Posiciones éstas consideradas estratégicas para captar la atención del lector. Por lo que ha de ponerse especial cuidado en su redacción.

Concluida la redacción del *informe,* siempre es aconsejable que otros profesionales lean el documento antes de su presentación o edición. Sus sugerencias siempre suponen mejoras.

La edición definitiva del *informe* va, con frecuencia, precedida de varios borradores y el empleo de muchas horas de trabajo, pues "la redacción de la investigación siempre implica mucha reescritura y refinamiento" (Gilbert, 1993: 334).

Lecturas complementarias

Alexander, C. (1986): *Writing for social scientists: how to start and finish your thesis, books or article.* Chicago, Chicago University Press.

Alvira, F. (1995): "Formas de presentación y representación gráfica de los datos. Preparación del informe". En MAP (ed.): *Sociología. Temas de materias sociales,* vol. 2, pp. 245-275.

Calvert, P. (1991): "Writing skills". En Allan, G. y C. Skinner (comp.) *Handbook for research students in the social sciences.* London, The Falmer Press, pp. 96-106.

Cooper, H. M. (1984): *The integrative research review: a systematic approach.* Beverly Hills, Sage.

Eco, U. (1982): *Cómo se hace una tesis.* Barcelona, Gedisa.

Gilbert, N. (1993): "Writing about social research". En Gilbert N. (comp.) *Researching social life.* Londres, Sage, pp.327-344.

Richlin, J. y Strenski E. (1991): *A guide to writing sociology papers.* New York, Saint Martin.

Sierra Bravo, R. (1986): *Tesis doctorales y trabajos de investigación.* Madrid, Paraninfo

Ejercicios propuestos

1. Elija dos investigaciones, una publicada en una revista científica y la otra como monografía. Compare la presentación que en ambas se hace de la investigación: la estructura y redacción del informe.
2. Elabore un proyecto de investigación sobre un tema de su interés. En él incluya todos los componentes esenciales de un proyecto y justifique el diseño de su investigación: por qué ha elegido unas estrategias y técnicas con preferencia a otras.